

esec

ESCOLA SUPERIOR DE EDUCAÇÃO



INSTITUTO POLITÉCNICO
DE COIMBRA

Departamento de Artes e Tecnologias

Mestrado em Human Computer Interaction

Relatório Final

Imersividade nos Videojogos e suas interfaces: Potencialidades do Reconhecimento de Voz

Daniela Sofia de Andrade Correia

Coimbra, 2017

esec

ESCOLA SUPERIOR DE EDUCAÇÃO



INSTITUTO POLITÉCNICO
DE COIMBRA

Daniela Sofia de Andrade Correia

Imersividade nos Videojogos e suas interfaces: Potencialidades
do Reconhecimento de Voz

Relatório Final de Mestrado em HCI, apresentado ao Departamento de
Artes e Tecnologias da Escola Superior de Educação de Coimbra para
obtenção do grau de Mestre

Constituição do júri:

Presidente: Professora Doutora Maria Fátima Neves

Arguente: Professora Doutora Rita Teixeira

Orientador: Professor Doutor João Orvalho

Trabalho realizado sob a orientação do Prof. Doutor João Orvalho

Junho de 2017

Agradecimentos

Gostaria de agradecer a todos os participantes que se voluntariaram a fazer os testes de usabilidade e a testar o reconhecimento de voz.

Um agradecimento especial a Mathieu Desautels e Miguel Fontoura que disponibilizaram o seu tempo para me ajudar e ensinar sobre o tema em questão deste relatório.

Imersividade nos Videojogos e suas interfaces: Potencialidades do Reconhecimento de Voz

Resumo: Este relatório documenta o estudo das potencialidades que a tecnologia de reconhecimento de voz tem quando utilizada nos videojogos, tanto a nível da interface do utilizador como a nível da imersividade, podendo comparar as vantagens e desvantagens desta tecnologia.

O objectivo desta pesquisa foi perceber porque é que o reconhecimento de voz não é mais utilizado na indústria de videojogos, visto que esta tecnologia não é recente e tem a capacidade de simplificar a utilização de um jogo, diminuindo o lixo visual ao manter a interface mais limpa e tornando o *gameplay* mais imersivo. Pretendo compreender como e se esta tecnologia consegue realmente melhorar os videojogos e a experiência dos jogadores, tendo por finalidade especificar as situações que aceitam o reconhecimento de voz como solução e aquelas em que esta tecnologia poderá piorar a experiência ao jogador.

Através de dois jogos e com a ajuda da participação de dez voluntários, foram feitos testes para entender as dificuldades e obstáculos que o utilizador sente ao utilizar a voz como interacção com o jogo e quais os pontos onde este sentiu melhoria.

Com estes testes e pesquisa, descobrimos, ao longo deste relatório, que o reconhecimento de voz é realmente uma tecnologia divertida e que ajuda a solucionar alguns problemas no desenvolvimento de um videojogo, caso este seja bem implementado, tendo criado uma lista de pontos-guia para ajudar a tornar o uso do reconhecimento de voz um sucesso nesta indústria de entretenimento.

Palavras-chave: Videojogos para computador, reconhecimento de voz, imersividade nos jogos.

Immersiveness in Video Games and its user interfaces: The Potential of Speech Recognition

Abstract: This paper documents the study of speech recognition and its potentials when used in video games in terms of user interface and immersiveness, allowing us to compare the advantages and disadvantages of this technology

The purpose of this paper is to understand why speech recognition is usually avoided in the gaming industry, as this technology is far from being new to us and has the ability to simplify the use of a game, reducing the visual distractions by keeping the interface clean and thus making its gameplay more immersive. Through this research I try to understand how and if this technology can actually enhance the video game and the experience of the players, my purpose being to specify the situations that accept speech recognition as a solution and those in which this technology could worsen the experience of the players.

Through various experiments with two games and the help of ten volunteers who tested them, I was able to observe and understand the difficulties and obstacles that one feels when using the voice as interaction with the game and which of its aspects were felt to be improved.

In this paper, throughout the research, I found that the voice recognition is really fun to use as an interaction technology in a video game if it is well implemented, having noted a list of guidelines to help make the most of this technology when developing a game as to make it a more probable success.

Keywords: Computer video games, speech recognition, immersiveness in games.

Índice

1. Introdução	13
1.1. Objectivos	17
1.2. Plano de Trabalho.....	18
2. Reconhecimento de voz.....	19
2.1. Aplicação do Reconhecimento de voz	22
2.1.1. Vantagens e Desvantagens	24
2.2. O Reconhecimento de voz nos videojogos.....	25
2.2.1. Videojogos com reconhecimento de voz	26
2.2.2. Ideias preconcebidas	28
2.2.3. Vantagens e Desvantagens em videojogos	29
3. Game User Research.....	31
4. HCI nos videojogos e a sua importância.....	39
5. Entrevista a <i>Game Developers</i>	43
5.1. Entrevista a Mathieu Desautels	45
5.2. Entrevista a Miguel Fontoura	49
5.3. Apanhado das entrevistas	52
6. Testes.....	53
6.1. Personas	56
6.2. Jogos escolhidos	59

6.3. Condições e procedimentos dos testes.....	60
6.4. Participantes.....	65
6.5. Hipóteses.....	67
6.6. Resultados.....	68
6.6.1 Resumo dos Resultados	75
6.7. Heurísticas de Nielsen aplicadas aos jogos testados.....	77
7. Conclusões.....	83
7.1. Guia para a implementação do reconhecimento de voz nos videojogos..	86
8. Bibliografia.....	89

Abreviaturas

HCI – Human Computer Interaction (Interacção Humano-Computador)

PC – Personal Computer (Computador Pessoal)

RPG – Role Playing Game (Jogo de Interpretação de Papéis)

NPC - Non-Player Character (Personagem que não é um jogador)

Conceitos e Definições

Human Computer Interaction (HCI) – Em português, Interação Humano-Computador, é uma matéria interdisciplinar que relaciona a ciência da computação, artes, design, ergonomia, psicologia, sociologia, semiótica, linguística, entre outras, e que estuda a interação entre pessoas e computadores, a qual é feita através da interface do utilizador, ajudando a torna-la o mais simples, seguro e agradável possível, maximizando a nossa capacidade de sucesso para com o sistema e minimizando o tempo desperdiçado que gastamos para o perceber e aprender a utilizar.

User Interface: Em português, Interface do utilizador. É o espaço, seja este físico ou virtual, onde a interação entre humanos e máquinas ocorre. É através da interface do utilizador que podemos interagir com o computador, controlando-o e operando-o, e também receber informação acerca do que se está a passar, alertando o utilizador e auxiliando-o na tomada de decisões.

Reconhecimento de Voz – Tecnologia que permite os computadores, através da utilização de um microfone, interpretar a fala humana, podendo ser utilizado, por exemplo, para transcrição ou como método de comando.

Game Developers – Em português, aquele que desenvolve jogos. São produtores de software que criam videojogos, podendo ser pessoas ou empresas.

Gameplay – Em português, Jogabilidade, é um termo na indústria de videojogos que inclui todas as experiências do jogador durante a sua interação com os sistemas de um jogo.

Sistema Operativo – Programa ou conjunto de programas cuja função é gerenciar os recursos do sistema, como definir qual programa recebe a atenção do processador, gerenciar memória, criar um sistema de arquivos, entre outros, fornecendo uma interface entre o computador e o usuário, e o qual é executado imediatamente após a máquina ser ligada. (Ex: Windows; iOS; Android...)

Role Playing Games – Em português, jogo de interpretação de papéis. É um tipo de jogo em que os jogadores assumem papéis de personagens e, através das suas escolhas ao longo do jogo, desenvolvem a história.

Non-Player-Character - Em português, uma personagem que não é um jogador. É uma personagem de qualquer jogo eletrônico que não é controlado por um jogador mas com o qual se pode interagir e que se envolve de alguma forma no enredo de um jogo.

Sandbox – Literalmente traduzido para “caixa de areia.” É um termo utilizado para definir os video-jogos nos quais o jogador foi liberto da estrutura tradicional que se encontra tipicamente nestes. Em vez de existir apenas uma direcção na qual o jogador pode ir, seja esta um caminho ou a nível de história, é dado ao jogador a capacidade de escolher o quê, quando e como eles querem abordar as situações e opções dadas no video-jogo.

Game User Research – Em português, pesquisa sobre o Utilizador de Jogos. Nesta pesquisa investigam-se formas de melhorar a experiência do utilizador de jogos e de certo modo pode-se dizer que é o estudo da noção de divertimento, tendo como objectivo tornar o jogo mais imersivo, mais intuitivo e mais acessível a todos os tipos de indivíduos e desta forma também alargar o seu público-alvo.

Heurísticas de Nielsen – Guia feito por Jacob Nielsen para a avaliação da interface do utilizador nos sistemas constituído por 10 pontos:

1. Visibilidade do sistema;
2. Correspondência entre o sistema e o mundo real;
3. Controle e liberdade do usuário;
4. Consistência e padrões;
5. Diagnóstico e recuperação de erros;
6. Prevenção de erros;

7.Reconhecimento ao invés de recordação;

8.Flexibilidade e eficiência no uso;

9.*Design* e estética minimalista;

10. Ajuda e documentação.

1. INTRODUÇÃO

1. Introdução

Os jogos são vistos, muitas vezes, como uma coisa fútil, infantil e às vezes inútil, no entanto, muitos não se apercebem da importância que estes têm no crescimento de um Ser Humano. É ignorado o facto de sempre ter existido o conceito de jogo em qualquer cultura de qualquer povo, podendo até ser encontrado no reino animal, servindo como uma simulação de situações reais, problemas, obstáculos e perigos, uma preparação divertida para as complicações da vida.

Com o desenvolver da tecnologia, também os jogos foram sofrendo uma transformação, criando os videojogos que tantos gostam. Neste mundo virtual, crianças e adultos aprendem mais sobre o mundo num ambiente onde podem experimentar sem medo de errar, divertindo-se e desafiando-se ao mesmo tempo, oferecendo-lhes as ferramentas para testar e exercitar as suas capacidades a nível cognitivo, cultural, social, gerencial, moral, lógico, criativo, entre muitas outras, criando jogos que promovem o convívio entre jogadores e a ajuda mútua, inculcando a importância do trabalho em equipa, colocando-lhes obstáculos ou puzzles que os fazem pensar e que oferecem recompensa ao serem resolvidos, instruindo na gestão de dinheiro e a saber poupar, melhorando a sua comunicação e até a aprenderem uma linguagem nova, ensinando a trabalhar para a obter realização pessoal e levando-os a querer completar os seus objectivos na vida real através de associação. Todos querem ser o protagonista da sua história e deixar uma marca.

Os videojogos conseguem levar uma pessoa a levantar-se e a querer ser alguém, a querer fazer algo ou parte de algo, no entanto, acaba por ser uma forma de arte que não recebe a devida atenção de muitos, sendo alvo de críticas e queixas muitas vezes injustas. Mesmo assim, a indústria de videojogos não deixa de evoluir constantemente. Os *Game Developers* tornaram-se mais audazes, mais competitivos, querendo surpreender sem terem medo de experimentar novos métodos e estilos de jogo, aprendendo com a comunidade que se formou à sua volta, querendo chegar a públicos mais diversos e a melhorar a experiência do jogador tendo em vista tornar

os seus produtos cada vez mais imersivos, fáceis de utilizar e compreender, mantendo o jogo limpo, atractivo e interessante.

Contudo, e apesar de esta tecnologia não ser novidade, só recentemente se começou a aceitar as potencialidades do reconhecimento de voz como método de interacção entre o utilizador e o jogo. Assim, descobri que ainda há uma falta de informação, estudo e pesquisa sobre esta tecnologia, não existindo um esclarecimento sobre a melhor forma de esta ser aplicada na indústria de videojogos, uma falha que eu tenciono ajudar a colmatar através deste relatório, focando-me nos videojogos para PC que utilizam linguagem inglesa ou fictícia.

Neste trabalho, irei identificar por que razão esta tecnologia se encontra sub-aproveitada, como ela poderá ajudar no desenvolvimento de jogos, aumentar e melhorar a experiência do utilizador e concluir em que casos ela deverá, ou não, ser utilizada.

1.1. Objectivos

Com esta pesquisa, pretendo compreender o motivo pelo qual o reconhecimento de voz enquanto modo de interactividade não é mais utilizado. Procuro saber qual a qualidade da tecnologia de reconhecimento de voz que existe hoje em dia e quais os seus limites e potencialidades dentro do desenvolvimento de videojogos para PC, focando-me apenas na língua inglesa e fictícia. Pretendo descobrir e compreender que tipo de videojogos beneficia com este modo de interacção e quais ficam prejudicados com ela.

Existem várias questões sobre o Reconhecimento de Voz em aberto, aquelas que eu irei abordar serão as seguintes:

- Ajuda a simplificar a *User Interface* do jogo, retirar informação a mais/lixo visual?
- Ajuda no acesso de opções e acções do jogo, permitindo maior complexidade de *gameplay*, quebrando o número limitado de teclas do teclado?
- Consegue tornar um jogo mais imersivo?
- Será prático para o utilizador?
- Em que situações deve ou não ser utilizado?

O meu objectivo é facilitar a resolução de alguns problemas a nível de desenvolvimento de jogos e do seu *Game Design*, assim como melhorar a experiência do utilizador enquanto jogador de videojogos, no seu *gameplay* e imersividade, através de um guia ou um conjunto de regras para a utilização deste tipo de tecnologia, uma lista de o que fazer e o que não fazer, direccionado aos *Game Developers*.

1.2. Plano de Trabalho

Primeiramente, irei estudar e comparar as informações, *feedback* e *reviews* já existentes sobre esta tecnologia e jogos para PC que a utilizem.

A abordagem que farei às questões que apresentei previamente será feita através da recolha, análise e comparação de dados, os quais serão obtidos através da pesquisa, entrevista a dois *game developers* e dos testes de usabilidade com utilizadores entre os 18 e os 30 anos de idade, os quais serão registados em vídeo, para observação de comportamento do indivíduo, juntamente com o método *think-aloud* e uma entrevista final. Os vários indivíduos jogarão dois videojogos que têm ou possibilitam a utilização de reconhecimento de voz, cujo *gameplay* será também gravado, a partir do qual será retirado o tempo que demora aos participantes a dizer os comandos e se o reconhecimento de voz detecta, entende e executa.

2. RECONHECIMENTO DE VOZ

2. Reconhecimento de voz

O reconhecimento de voz é uma tecnologia que permite os computadores conseguirem interpretar os sons e a fala humana através da utilização de um microfone.

Esta tecnologia trás novas possibilidades à sociedade na forma de melhorias, forma de quebrar limitações ou chegar a um novo tipo de público, modificando métodos de trabalho, interacções com aparelhos e sistemas, entre outras coisas fazendo com que esta tecnologia seja imensamente importante de ser estudada e desenvolvida [24].

2.1. Aplicação do Reconhecimento de voz

A tecnologia de reconhecimento de voz é utilizada em muitos e diferentes campos, muitos deles situações e objectos com que nos deparamos no dia-a-dia. Podemos ver esta tecnologia utilizada em veículos, telemóveis, electrodomésticos, casas inteligentes, entre outros [24].

Dou os seguintes exemplos da sua utilização e algumas das suas vantagens:

A utilização de reconhecimento de voz num veículo pode tornar mais seguro a utilização de certas funcionalidades enquanto se guia. Esta tecnologia deixa o condutor utilizar comandos simples sem perturbar a sua condução, por exemplo, conseguir mudar de rádio ou de música, fazer chamadas ao ter o telemóvel conectado ao sistema do carro, etc...

Na área de saúde faz-se uso desta tecnologia ao possibilitar os médicos a fazerem relatórios rápidos durante operações e análises, falando alto enquanto o computador interpreta e escreve o que foi dito. Também é utilizada como método de ajuda na terapia da fala e na assistência a pacientes com problemas a nível motor, tanto dentro como fora dos hospitais.

O uso de reconhecimento de voz a nível militar também é comum, incluindo aplicações nos seus aviões de caça para mudar a frequência de rádio, comandar o sistema de piloto automático, definir coordenadas e parâmetros de lançamento de armas e controle de voo através da voz. O exército, e não só, utiliza esta tecnologia para algumas simulações e treinos, como por exemplo, fazer a formação de controladores de tráfego aéreo, onde o sistema reconhece os comandos do formando e simula de acordo.

Esta tecnologia é também importante na educação, para aprender uma segunda língua, para ensinar crianças cegas que não conseguem escrever ou ver bem onde estão as coisas num ecrã e podem usar a voz para executar comandos, para ensinar crianças com problemas motores não se precisam de preocupar com escrever, etc...

Existem imensas funções para o reconhecimento de voz e demoraria imenso falar sobre todas elas. Entre todas as possibilidades de aplicação desta tecnologia,

estão também: A possibilidade de dar comandos por voz nos videojogos; Casas Inteligêntes, ligando a luz ou funcionando com electrodomésticos através da voz; Atalhos por voz para o telemovel ou para computadores e subtítulos automáticos para vídeos.

2.1.1. Vantagens e Desvantagens

Vantagens:

- Deficiências: Uma das maiores vantagens desta tecnologia é a sua capacidade de tornar um *software* apto para ser utilizado por indivíduos com deficiências, principalmente se estas afectarem o corpo a nível motor.
- Rapidez: Torna-se bastante mais rápido pronunciar apenas uma palavra ou simples frase para executar um comando ou até ditar um texto para o computador escrever do que ser o próprio utilizador a escrever.
- Multitarefa: Um ponto atractivo desta tecnologia é a possibilidade de dar comandos por voz ao mesmo tempo que se tem as mãos ocupadas a fazer outra acção, por exemplo, a usar teclado e rato ou a guiar um carro, sem que estas se atrapalhem uma à outra.

Desvantagens:

- Barulho: O uso desta tecnologia não é nada discreta e nem toda a gente quer ser alvo de atenção por a usar, para além de que não se pode fazer barulho em qualquer lugar ou a qualquer hora, por exemplo, jogar um jogo à noite sem querer incomodar quem está a dormir.
- Vocabulário Limitado: Apesar de ser necessário apenas uma palavra ou pequena frase para executar o comando, esta tem que ser pronunciada com palavras específicas as quais o indivíduo tem que memorizar, podendo não ser familiar ou cómodo para o utilizador.
- Delays e Interpretações erradas: Apesar dos programas que usam a tecnologia de reconhecimento de voz serem desenvolvidas para serem rápidas e interpretarem a fala do utilizador, nada impede que não hajam erros ou falhas nos piores momentos.

2.2. O Reconhecimento de voz nos Videojogos

Esta tecnologia possibilita um novo tipo de interação nos videojogos, deixando o jogador utilizar a voz para dar comandos e ordens, acrescentando mais uma opção ao modo de controlo do jogo e afectando o estilo de *gameplay* deste.

Existem vários motivos que atraem a exploração do reconhecimento de voz nos videojogos. Uma delas é o facto de esta tecnologia não requerer *hardware* complicado nem caro e de os sistemas operativos recentes incluírem *software* gratuito que suporta o reconhecimento de voz, o que faz com que esta tecnologia seja bastante acessível a toda a comunidade, ao contrário de tecnologias como as de reconhecimento de movimento (Ex: Kinect), *eye-tracking*, tecnologias de reconhecimento de ondas cerebrais (Ex: EMOTIV), capacete ou óculos de realidade virtual (Ex: Oculus Rift), entre outras. Outro motivo que o torna interessante é o facto de o reconhecimento de voz ter vastas possibilidades a nível de uma interface mais limpa e mais fácil de compreender e uma maior imersividade e acessibilidade nos videojogos, ou seja, esta tecnologia de mãos livres pode melhorar esta forma de entretenimento não só de um modo geral mas também expandindo a indústria de videojogos a um público diferente. (Ex: Indivíduos com problemas a nível motor.)

Como a linguagem humana possibilita um número quase ilimitado de sons, palavras e frases, as quais podem ser mapeadas pelo sistema, o reconhecimento de voz pode tornar-se benéfico para videojogos que necessitem ou que requerem inúmeros comandos e/ou inúmeras combinações de teclas. É de mencionar, que a capacidade para um indivíduo se lembrar e recordar comandos numa linguagem natural é muito maior quando comparada com lembrar e recordar combinações de teclas arbitrárias, tornando esta tarefa mais simples e familiar.

Claro que, esta tecnologia, não deve ser simplesmente implementada num videojogo e pronto. Os *designers* e *developers* de jogos precisam de “...entender os benefícios e limitações desta tecnologia. Entender quando faz sentido utiliza-la e quando esta poderá simplesmente causar problemas.” [24].

2.2.1 Videojogos com reconhecimento de voz

Tal como mencionei anteriormente, a utilização de reconhecimento de voz na indústria dos videojogos não é algo novo, existindo alguns títulos cuja interação é feita através desta tecnologia e outros nos quais os jogadores acrescentam modificações que possibilitam o uso de voz. Existem também programas de voz que podem ser customizados pelo jogador e corridos em conjunto com o videojogo, associando um som, palavra ou frase a uma tecla.

Dois jogos conhecidos que tem reconhecimento de voz são o “*Mass Effect 3*” da BioWare [1], um *RPG* de ficção científica em que o jogador desempenha o papel de um comandante, o qual tem de dar ordens à sua equipa de NPCs durante o jogo, tendo a possibilidade de usar a voz para o fazer, e “*Tom Clancy’s Endwar*” da Ubisoft [2], um jogo de estratégia em tempo real e simulação militar, no qual, durante as batalhas, a utilização de voz é crucial para informar, alertar e comandar a nossa equipa de NPCs. Tanto um como o outro utilizam o reconhecimento de voz para o mesmo fim, para dar ordens às unidades da tua equipa durante o combate. Em ambos estes jogos, o jogador pode dar ordens por voz durante os combates às outras personagens, desocupando as mãos do jogador para mexer a sua personagens, apontar, disparar, etc... ajudando o jogador a fazer acções em simultâneo com maior facilidade. Como a acção destes jogos não é propriamente de ritmo acelerado, focando-se na estratégia e dando tempo ao jogador para preparar e modificar a sua tática, o tempo de utilização dos comandos por voz não condiciona a jogabilidade. A maior parte da comunidade de jogadores elogiou bastante a utilização desta tecnologia, dizendo que tornava a sua experiência diferente, mais interessante e divertida, no entanto uma pequena minoria não se sentiu agradada, queixando-se que tornava-se aborrecido estar constantemente a repetir comandos e que preferiam a utilização do teclado. Uma das queixas direccionadas ao vídeojogo “*Tom Clancy’s Endwar*” é sobre o seu sistema de reconhecimento de voz levantando a questão da capacidade, qualidade e limitações que existe nesta tecnologia, queixando-se que os comandos por voz funcionavam melhor na consola do que no PC.

Um jogo bem conhecido no qual se pode aplicar uma modificação para a utilização de reconhecimento de voz é o “*The Elder Scrolls: Skyrim*” da Bethesda[3], um *sandbox* RPG de fantasia no qual o jogador tem o papel de um aventureiro, explorador e herói. Esta modificação que foi muito bem recebida pela comunidade, dando a possibilidade aos jogadores de utilizar a voz para produzir um poder mágico especial chamada “*Shout*”(grito) ao exclamar, para o microfone, as palavras fictícias certas para produzir o “*shout*” que querem utilizar. Assim, quanto se utiliza uma mão no teclado para o movimento, a outra no rato para controlar as mãos do personagem (atacar, defender, etc...), o jogador utiliza a voz como a própria voz da sua personagem, tornando o comando bastante natural e imersivo. Como este poder não é utilizado em todas as situações, não pode ser repetido em demasia e nos momentos em que é utilizado não há pressão a nível de tempo, o reconhecimento de voz neste jogo é fácil de usar e não se torna chato.

Existem vários programas que possibilitam usar voz para dar ordens e usar comandos em videojogos que não estavam previamente preparados para tal, por exemplo o *Tazti* [4], o *VoiceBot* [5], entre outros. Nestes programas, o utilizador tem, normalmente, a possibilidade de criar perfis onde o jogador grava as várias configurações. Estes programas dão a possibilidade de uma maior flexibilidade desta tecnologia, podendo solucionar alguns problemas como a utilização de palavras ou frases mais familiares para o utilizador do que aquelas que o videojogo obriga a usar e a situação conhecida do sistema não reconhecer bem o comando do jogador devido ao sotaque ou tom de voz deste.

2.2.2. Ideias preconcebidas

A utilização do reconhecimento de voz nos videojogos possibilita

- Uma interface mais limpa e perceptível;
- Uma maior imersividade;
- Maior facilidade a controlar o jogo;
- Quando customizável, uma maior liberdade e uma adaptação mais fácil ao jogo;
- Possibilita mais acções em simultâneo.

2.2.3 Vantagens e Desvantagens em videojogos

Para além das vantagens e desvantagens já mencionadas anteriormente sobre o reconhecimento de voz no geral, as vantagens e desvantagens desta tecnologia quando aplicadas especificamente nos jogos são:

Vantagens:

- **Memorização:** Não existe necessidade de memorizar controlos e botões, sendo apenas preciso pronunciar palavras ou frases simples já familiares para o jogador.
- **Aumento de Comandos:** O jogador não está limitado ao número de botões disponíveis. Caso o jogo dê a possibilidade de o jogador executar bastantes acções, pode-se utilizar botões simples e voz, em vez de se adicionar combinações complicadas de teclas.

Desvantagens:

- **Interferência na Comunicação:** Caso o jogo seja *online* e necessite do uso de voz para comunicar com, por exemplo, outros membros da equipa, o uso da voz para dar comandos poderá interferir com a fala normal e vice-versa.
- **Momentos Críticos:** A maior parte dos jogos são feitos de modo a que o jogador necessite de algumas habilidades para passar alguns obstáculos, sendo um deles a rapidez e precisão de tempo. Tal como mencionei, nada impede que não hajam erros ou falhas na tecnologia de reconhecimento de voz e tal pode acontecer num momento crítico, fazendo o jogador falhar e ficar frustrado.

3. GAME USER RESEARCH

3. Game User Research

Game User Research, pode ser traduzido por, Pesquisa sobre o Utilizador de Jogos, é uma área extremamente recente na indústria de videojogos e dedica-se à análise, desenvolvimento e melhoria da experiência do jogador.

À medida que se avança no tempo, as pessoas ficam cada vez mais informadas sobre a importância do desenvolvimento da indústria de jogos e, em paralelo, cada vez mais as companhias de videojogos incluindo as grandes e as mais importantes desta área, encetam esforços tendo em vista dar resposta à constante necessidade de evolução dos mesmos, para isso vão recorrendo a uma variedade de técnicas de HCI e UX, para assim poderem analisar a eficácia do design e gameplay na experiência do jogador, isto durante todo o processo de melhoramento dos videojogos [15].

Várias companhias, como a Player Research e Serco ExperienceLab (Inglaterra), a Valve Software, Electronic Arts, BoltPeters e VMC Labs (EUA e Canadá), começaram já a usar a biometria (estudo estatístico das características físicas ou comportamentais dos seres vivos) para medir cientificamente a relação entre eventos ocorridos dentro de um videojogo e as emoções, sentimentos e reflexos da pessoa que o está a jogar. [15]

Sendo que o estudo da experiência do utilizador na área dos jogos é já hoje reconhecido e destacado publicamente pela sociedade do ponto de vista profissional, como exemplo disso temos a conferência da GDC, Game Developers Conference (Conferência de Desenvolvedores de Jogos) [23, 15] onde todos os anos, se reúnem profissionais da área de videojogos tendo aí lugar várias iniciativas, como exposições, eventos sociais, divulgação, tutoriais, workshops, atribuição de prémios entre outras, de modo a promover divulgar, inspirar e educar sobre este assunto.

O principal motivo que incentiva a evolução na área dos videojogos, a nível de design e interacção, é a necessidade de chegar de forma mais eficaz a mais audiência, resultando no aumento potencial de venda dos seus produtos.

Convém aqui salientar que uma das razões pela qual assistimos ao progresso nesta área é a necessidade de satisfazer ou de colmatar a falha na interface existente, na maior parte das vezes o estudo sobre melhorias nos videojogos, em geral, advém de problemas detectados numa audiência restrita, como seja; os cegos, os daltónicos, indivíduos com problemas motores, etc. No entanto, ao estudar-se a maneira de chegar a estas audiências consideradas “minorias”, também se progride a nível de interface e gameplay em relação às audiências olhadas como “maiorias”, resultando na inserção de novas possibilidades, soluções e facilidades, tendo em vista uma experiência melhor, mais divertida e até mais realista no jogo.

Apesar disto, a evolução em relação a HCI e UX é relativamente lenta nesta área, e muito do que vemos nos videojogos mais recentes pode também ser visto naqueles que são mais antigos, pois a interface, e o modo de interagir pouco se altera de uns para os outros [15]. Especificamente no que diz respeito ao estudo do reconhecimento de voz na indústria dos videojogos, apercebi-me de que não existe muita investigação, nem informação disponível ao público em geral e foi então este um dos motivos preponderantes que me levou a interessar-me e a debruçar-me sobre este tema para desenvolver a minha pesquisa.

Perceber o modo como o jogador experiencia um videojogo é extremamente importante para o desenvolvimento de videojogos; É aqui que entra a pesquisa do utilizador de jogos (Game User Research), ciência que estuda o indivíduo enquanto utilizador de videojogos. Esta ciência estuda a noção de divertimento do jogador, no entanto, o conceito de divertimento é relativo pois pode mudar de indivíduo para indivíduo, de lugar para lugar, de geração para geração, tornando esta área demasiado vasta. Assim, tenta-se entender os factores que afectam a diversão do jogador, sendo estes, a imersão no jogo ou seja o envolvimento emocional e cognitivo e a dissociação com o mundo real, o desafio do jogo e a fluidez deste que é a noção de controlo, distorção da noção de tempo e perda de consciência de si próprio. [15, 19, 20].

Ora este tipo de pesquisa é extremamente recente e há ainda uma grande falta de investigação nesta área, não só a nível de o que é explorado e estudado mas também nos meios e métodos utilizados para o fazer com mais eficácia e veracidade [15, 18, 20, 21].

Este estudo envolve uma grande quantidade de dados, que muitas vezes estão dispersos e misturados resultando numa difícil compreensão [21], pois não só as interações do jogo como o próprio jogador, um ser humano, é demasiado complexo para uma leitura fácil e análises simples, algo que é referido, cito “Isto apenas nos lembra de que as experiências humanas são realmente muito complexas...” [18]. Assim sendo, os métodos utilizados requerem a ajuda de aparelhos e software para conseguirem recolher dados, analisá-los e depois trabalhá-los. Estes métodos podem ser auto-relatórios (Ex: entrevistas, questionários, think-aloud), respostas fisiológicas (Ex: pulsação, direcção do olhar), registos de actividade (Ex: telemetria, duração de tarefas) e psicofisiologia, ou seja, estados psicológicos deduzidos de respostas fisiológicas (Ex: Pupila dilatada, temperatura do corpo e expressões faciais) [21].

No entanto, estes métodos nem sempre conseguem dar toda a informação importante e necessária quando utilizados por si só e podem também afectar negativamente o próprio estudo e nem sempre há as condições necessárias para um teste onde os dados estejam completamente imperturbados. A situação do jogador não estar tão confortável num laboratório de testes como estaria em sua casa ou o humor do participante enquanto executa o teste pode afectar os dados. [18, 20]. Aliás, nem sempre é possível utilizar certos aparelhos que seriam de grande ajuda a recolher dados concretos, podendo estes ser muito caros, necessitar de um profissional para mexer nele ou até por serem demasiado evasivos e afectarem a veracidade dos testes, tanto na recolha de informação como na análise e organização da mesma, requerendo muitas vezes softwares também específicos. Às vezes há necessidade destes softwares serem desenvolvidos e programados de base para que possam atender às necessidades de um estudo específico [20].

Assim, a tarefa de estudar um jogador não é tão trivial como parece, havendo imensos factores a afectar a eficácia destes, muitos deles devido à ignorância e falta de pesquisa de quais métodos devem ser usados e em quais situações [18, 20].

Mesmo assim, existem alguns guias ou exemplos de testes pelos quais nos podemos basear. Os métodos tradicionais usados na área de Game User Research são preferivelmente os qualitativos, estes incluem recolher dados subjectivos através de Think-Aloud (O acto de verbalizar o que se sente, pensa e experiencia), entrevistas e observações directas do utilizador, devido a estes métodos terem resultados razoavelmente precisos [18, 20]. No entanto, estes métodos tem as suas desvantagens, o utilizador estudado e entrevistado pode sentir-se observado e não se sentir confortável o suficiente para agir ou falar naturalmente. Pedir ao jogador para verbalizar o que sente e pensa pode distrair-lo e perturbar a sua experiência no jogo, para além de que é difícil uma pessoa tentar articular as suas experiências complexas oralmente enquanto têm que se focar no jogo. A utilização de câmaras, mostrando depois o vídeo ao jogador para que este possa fazer uma introspecção pode solucionar alguns destes problemas, no entanto, este pode não se lembrar de todos os detalhes ou situações importantes a relatar [18, 20].

Por outro lado, métodos quantitativos, tem o potencial de recolher dados mais precisos, verdadeiros e diversos de uma forma activa ao longo do teste. Os métodos quantitativos utilizados actualmente são, normalmente, a telemetria e a psicofisiologia [20].

A telemetria lida com o jogador dentro do jogo, ou seja, a observação das decisões do personagem do jogador, ou seja, acções, movimentos, hesitações. A vantagem deste tipo de observação é que é possível capturar objectivamente o modo de jogo sem distrair o jogador, podendo até ser feita não-presencialmente. No entanto, os dados estão limitados ao que se consegue observar neste mundo virtual e não captura a experiência do jogador no seu todo, particularmente, o modo como este sente [20].

Recentemente, a psicofisiologia emergiu como um modo de quantificar a experiência do jogador. Estes métodos deduzem estados psicológicos através de respostas fisiológicas detectadas utilizando medidas como a frequência cardíaca, electrocardiograma, electromiografia, electroencefalografia, temperatura do corpo e dilatação das pupilas, entre outros, sendo os métodos mais populares a detecção e medição da variação elétrica da pele e eletromiografia facial, seguida da frequência cardíaca. Tal como a telemetria, estes métodos que detectam respostas fisiológicas conseguem capturar esses dados continuamente, em tempo real e com precisão [20].

Infelizmente, e apesar dos investigadores desta área darem muito valor à importância dos dados fisiológicos, pois vêem-nos como sendo mais objectivos de natureza e feedback mais credível, estes métodos têm falta de contexto para uma interpretação correcta. É por isso que, a maior parte das vezes, os investigadores têm alguma preferência em utilizar o Think-aloud em par com os outros métodos [18].

Existe, tal como mencionei anteriormente, uma grande falta de conhecimento a nível de que métodos usar, quando e como. Assim, existe também uma falha no entendimento de como se deve organizar, comparar e combinar dados quantitativos e qualitativos para avaliar as experiências dos jogadores, principalmente quando esses dados vem de mais do que um método de naturezas diferentes. No entanto, apesar das dificuldades e no tempo gasto em demasia para fazer este tipo de combinação, devido a alguma falta de conhecimento de como o fazer, misturar métodos prova ser bastante eficiente obter dados e introspeções úteis no que diz respeito à experiência do jogador e a momentos/eventos de maior interesse no jogo e no gameplay [18].

4. HCI NOS VIDEOJOGOS E A SUA IMPORTÂNCIA

4. HCI nos videojogos e a sua importância

O *design* dos videojogos pode afectar a percepção e o comportamento do utilizador e, como este é uma forma de *software* interactivo, é necessário estudá-lo pela perspectiva do HCI de modo a conseguirmos aperfeiçoar o seu *design*, melhorando a sua eficiência, minimizar os seus erros, tornar a aprendizagem e a sua utilização mais fácil, tentando chegar a um público o mais diverso possível, e aumentar a satisfação do utilizador.

Antigamente, a pesquisa de HCI em videojogos raramente os considerava diferentes de outras formas de *software* e isto levava a que a investigação não abordasse os videojogos como jogos especificamente, conseguindo apenas, com estes estudos, retirar informação e levantar questões que, apesar de úteis, exploram apenas o básico no que diz respeito ao *design* de interfaces nestas formas de entretenimento.

É preciso compreender que a interacção encontrada nos videojogos é bastante distinta da interacção encontrada noutros tipos de *software* e que a motivação para os usar é diferente da motivação para jogar um jogo. Ao contrário da maior parte dos programas e aplicações, os videojogos não são feitos para suportar actividades externas definidas pelo utilizador. Em vez disso, definem as suas próprias actividades nas quais os jogadores participam, ou seja, os videojogos contêm um sistema de valores que os jogadores percebem e adoptam, moldando o jogo.

Devido a estas diferenças, os videojogos devem ser avaliados com outro tipo de cuidados. Existem videojogos que são difíceis de aprender, pedem soluções ineficazes aos desafios apresentados, desafiam a memória do jogador e empurram-no a cometer erros, no entanto tudo isto é feito em nome da diversão, do desafio e competição. Obviamente, um outro tipo de *software* não deve obrigar o utilizador a ter que memorizar termos e controlos, nem a induzi-lo em erro, os interfaces destes programas e aplicações tendo sido desenvolvidos para serem eficazes e eficientes, podendo então ser avaliados como maus se o utilizador cometer demasiadas falhas, ficar confuso ou frustrado durante o uso do sistema, sendo mais fácil de identificar o que é um erro ou um defeito durante um teste experimental do que nos videojogos. Por exemplo, podemos considerar um erro de *gameplay* quando um jogador mata um

civil durante um videojogo? A resposta a essa questão depende muito nos valores do jogador e as razões e intenções deste.

Assim, as 10 Heurísticas de Nielsen [25, 26] não devem ser aplicadas aos jogos da mesma maneira que a outro tipo de *software*, como por exemplo, de acordo com as heurísticas, um software não deve nunca ser frustrante de ser utilizado pelo utilizador, no entanto, um videojogo pode querer utilizar essa frustração como motivação. De entre as dez heurísticas, todas elas importantes no desenvolvimento de videojogos, principalmente na interface, existe um ponto que acaba por sobressair mais nesta forma de entretenimento e é a partir dela que se consegue avaliar bem a qualidade do videojogo: A satisfação do utilizador [27].

5. ENTREVISTA A *GAME DEVELOPERS*

5.1. Entrevista a Mathieu Desautels

Tive a possibilidade de falar com Mathieu Desautels, não só um grande jogador de videojogos mas também um *software* e *game developer* da Warner Bros. Games (Montreal, Canada), que me deu um pouco do seu tempo para falar e discutir sobre o reconhecimento de voz nos videojogos, falando apenas por si e não pela companhia onde trabalha.

A pergunta que fiz que abriu a entrevista foi: Na tua experiência como jogador e como *developer*, qual a tua opinião sobre a tecnologia de reconhecimento de voz na área dos videojogos como método de imergir os jogadores um pouco mais no ambiente, história e situações do jogo e a manter a interface visualmente mais limpa e mais perceptível?

Mathieu responde: “Não existe qualquer dúvida que esta tecnologia pode trazer melhorias a nível de interfaces de videojogos. Ao se utilizar o reconhecimento de voz, certos botões podem tornar-se totalmente desnecessários e devem ser retirados do ecrã. Por exemplo, há uns tempos atrás adorava jogar um videojogo de *poker* que tinha muitos poucos botões... aliás, durante o jogo em si devia ter apenas um botão. Era quase tudo controlado pela voz!

Relativamente á imersão, na minha perspectiva enquanto jogador, quando jogo um videojogo com esta tecnologia sinto-me estúpido. Estou no meu quarto a falar sozinho feito louco e este sentimento desencoraja-me a voltar a jogar este tipo de coisas. Se calhar é culpa minha por não estar habituado, mas pessoalmente não gosto muito. Por outro lado, pode ser culpa de quem criou o jogo que não fez com que a utilização desta tecnologia fosse natural. Aliás, agora que penso nisso, o problema deve ser mesmo esse. Acabo de me lembrar que realmente já joguei jogos com reconhecimento de voz que adorei. Senti-me mesmo como se estivesse lá e o uso da voz parecia uma reacção óbvia aos acontecimentos do jogo. Não preciso de ir muito longe, olha, o jogo de *poker* que mencionei pedia que os jogadores dissessem em voz alta o que estavam a fazer, usando termos já conhecidos deste tipo de jogo, tal como se estivessemos numa mesa real com outros jogadores reais. Claro que, neste caso, os termos eram conhecidos e o jogador sabia exactamente que palavras

dizer para que o sistema do videojogo o compreendesse e isto ajuda a tornar toda a interacção mais natural, algo que não é tão normal noutros jogos onde o jogador tem que decorar frases...compostas de forma pouco comum. Frases que não dizemos habitualmente na rua, uns com os outros e assim. As pessoas não sabem bem o que dizer nem como falar com o sistema. Não é uma coisa imediata, é preciso tempo para aprender. Para tal, ou o jogo consegue fazer um reconhecimento de voz bastante familiar ou ir dando, pouco a pouco, direcções para que o jogador perceba o que tem que fazer e tenha tempo para se habituar ao modo de *gameplay* do jogo e ao seu sistema.”

Voltei então a questionar: E como *developer*?

Mathieu riu-se: “Como *developer*? Este tipo de tecnologia é uma dor de cabeça para se implementar! Entende, para que esta tecnologia funcione os *developers* tem normalmente que tentar adivinhar as intenções e reacções que os jogadores puderam ter para que o reconhecimento de voz funcione o melhor possível sem que tenhamos que preparar o jogo para a linguagem inglesa inteirinha com todas as formas e combinações possíveis. Isso era de loucos! É por isso que videojogos que tenham termos já conhecidos ou onde as respostas sejam simples, como sim ou não, são mais fáceis tanto para os que desenvolvem como para o jogador. É algo simples, sabe-se logo as opções existentes e o que dizer para confirmar ou negar algo.

Agora, quando os comandos são mais complicados, a memória do jogador pode falhar e a gramática tende a atrapalhar. Por exemplo, eu sei que quero que a personagem X vá para ali e existe uma ordem para isso. Mas será que digo: Personagem X, move-te para o Ponto A! Ou será que digo: Quero que vás para o ponto A, Personagem X! Se calhar é outro termo qualquer! É preciso ter cuidado com a maneira como os comandos são escolhidos para o jogador pronunciar, tem de ser bem pensado para que não haja confusões e seja o mais parecido com a linguagem familiar.

Mas este não é o único problema na implementação desta tecnologia! Olha, o sotaque. Este pode afectar o reconhecimento de voz muito facilmente em algumas situações.”

Ouvindo isto, intrevi: Há jogos que utilizam uma linguagem real, como o inglês, mas também existe a linguagem fictícia. O sotaque afecta o reconhecimento de voz em ambos os casos?

Mathieu hesita antes de responder: “Pode afectar... Quer dizer, realmente a linguagem fictícia pode ajudar bastante a nível de reconhecimento de um jogador que tenha sotaque. Obriga o jogador a dizer uma palavra específica para acontecer algo e como a linguagem é mais limitada, o que significa menos variações para o sistema, é mais fácil esta tecnologia perceber e interpretar o que foi dito. É claro que isto cria outro problema, a memorização de palavras estranhas. O que pode nem ser um problema se for bem feito! Olha, eu já não vejo os filmes do *Harry Potter* há muito tempo, mas é difícil esquecer-me de como se pronuncia alguns dos seus feitiços, como *Avada Kadavra* e *Wingardium Leviosa*! Estes são os mais conhecidos, porque lembro-me de palavreados estranhos de outros filmes e de outros jogos! O que quero dizer com isto, é que existem coisas que ficam no ouvido e há a possibilidade de conjugar isso, tornando fácil a memorização de comandos feitos através da pronunciação da linguagem fictícia.”

Com isto, coloquei uma nova questão: E quando à customização dos comandos de voz? Tal possibilidade não tornará a utilização do reconhecimento de voz mais fácil para os jogadores?

Mathieu explica: “Entende-se porque poderão pensar isso e a nível teórico até faz sentido... mas a nível prático não resulta nada bem! Há uma coisa que as pessoas não tem noção: Os jogadores são preguiçosos. Somos nós, enquanto *developers*, que temos de pensar nos problemas e descobrir como dar a papinha feita ao jogador.

Antigamente, os jogadores as coisas sozinhos. Eu não me lembro dos primeiros jogos do *Super Mário* darem pistas aos jogadores de que não se podia tocar em certos inimigos, que se podia saltar em cima de outros, o que é que os bonus, como o cogumelo, faziam... ou até que havia a possibilidade de se descer pelos tubos ao se clicar para baixo! Mas os jogadores descobriam as coisas sozinhos ao jogar e experimentar. Os jogadores eram curiosos!

Hoje em dia, se apresentamos uma coisa parecida, eles resmungão porque nós não os avisámos que aquilo era assim e que dava para fazer! Temos que ter tudo escrito e demonstrado porque a maior parte simplesmente não experimenta nem testa coisas novas ou diferentes do habitual e acha que não o tem que fazer!

Voltando então a questão de customização de voz... sim, isto dá possibilidades engraçadas ao jogador. Mas achas que a maior parte vai querer dar-se ao trabalho de perder tempo a preparar o sistema antes de jogar? Aliás, até podem estragar o jogo ao utilizar esta possibilidade de maneira errada! Por exemplo, customizar o reconhecimento de voz com palavras demasiado parecidas ou algo assim e criar erros de entendimento para o sistema! Os jogadores querem é o sistema implementado logo e ponto a usar.”

Como o tempo disponível estava quase a acabar, Mathieu procedeu a fazer uma pequena conclusão: “O reconhecimento de voz tem imensas possibilidades pouco exploradas e sim, que possibilitam uma maior imersão nos videojogos, dependendo imenso do modo como é implementado, pois pode melhorar ou arruiná-lo completamente! O *developer* deve ter em conta do tipo de videojogo que é, se realmente necessita ou não de reconhecimento de voz e porquê, se esta tecnologia está a trazer algum extra ou a facilitar algo ou a tornar o gameplay mais natural ou não e em que situações no jogo é que o jogador vai utilizar voz...se faz sentido e se dá jeito. Há que pensar bem na coisa e planear. Quer dizer, não vamos andar a implementar esta tecnologia em tudo quanto é sitio apenas por existir, não é? Não, há muitas variantes em que pensar. O ritmo do jogo, situações críticas em que algo pode falhar... claro, temos que nos lembrar que nada é infalível e qualquer sistema pode ter um delay ou uma falha de entendimento.

Há que entender a tecnologia e o jogo e saber evitar e resolver problemas. Acredito que se o *developer* souber o que está a fazer, poderá fazer coisas brilhantes com esta tecnologia na indústria dos videojogos.”

5.2. Entrevista a Miguel Fontoura

Tive também a oportunidade de falar com Miguel Fontoura, cuja paixão enquanto jogador o levou a entrar na indústria dos videojogos, sendo hoje game developer na *Collide* (Lisboa, Portugal). Apesar do tempo limitado que tive para falar com ele, consegui ter uma discussão interessante sobre o reconhecimento de voz nos videojogos, o qual fala apenas por si e não pela companhia onde trabalha, e tive a possibilidade de ouvir a sua opinião sobre o assunto.

Fiz-lhe então a pergunta: Na tua experiência como jogador e como *developer*, qual a tua opinião sobre a tecnologia de reconhecimento de voz na área dos videojogos como método de imergir os jogadores um pouco mais no ambiente, história e situações do jogo e a manter a interface visualmente mais limpa e mais perceptível?

Imediatamente, Miguel respondeu “É giro!”, rindo-se antes de acrescentar que “A tecnologia de reconhecimento de voz não é muito usada nos videojogos e raramente me deparo com jogos onde a tenha de usar para conseguir jogar. No entanto, quando a tive de usar, diverti-me e gostei da experiência.” Após uma pausa, disse “...mas pensando bem, não me lembro de grandes nomes de videojogos onde se use o reconhecimento de voz.”

Dei então o exemplo do “*Mass Effect 2*”, o qual o apanhou de surpresa, pois tinha-o jogado bastante e nunca tinha reparado, comentando que “É esse o problema de muitos jogos que têm implementado este tipo de tecnologia. Não está á vista e os jogadores, que não estão habituados a que exista a possibilidade de poderem usar voz, nunca vão adivinhar nem procurar. Convém que exista algum tipo de aviso... uma pista de que o podem fazer.” Miguel afirma: “Falar com uma máquina não é natural. Não é algo que um utilizador faça intuitivamente, ainda não. Convém que haja uma introdução, um pequeno alerta, assim, talvez o jogador adira a tal modo de jogar.”

Miguel continua: “Existem bastantes variáveis na qual uma pessoa tem de pensar para implementar este tipo de tecnologia. Uma pergunta que faria é, será que toda a gente tem microfone? Se calhar, hoje em dia, sim. Mas é uma questão a considerar, talvez dando a possibilidade de poder usar outro tipo de *input* para além da voz. Depois há a questão de... Faz sentido? Fazer com que o uso da voz pareça fazer sentido, que estamos a... falar com alguém ou a comandar algo, fazendo-nos esquecer que estamos a falar para uma máquina, tornando a fala natural, tal é possível! Se for bem feito. À que implementar o reconhecimento de voz quando é preciso e quando faz sentido, não é para se pôr de qualquer maneira só para dizer que tem a tecnologia ou só para retirar botões e informação do ecrã que seria melhor sendo visual.” Miguel dá o exemplo do jogo “*Dead Space*”, o qual substitui a típica barra de vida que vemos a um canto do ecrã por algo mais imersivo, fazendo a parte da armadura futurista da personagem, ao longo da coluna vertebral, ter uma luz que enche ou esvazia dependendo da vida da personagem. “É um sítio para onde um jogador olha muito e, apesar de ser fora do normal, sabe-se logo o que é e para que serve. Retira lixo visual ao mesmo tempo que informa e parece mais imersivo ter uma armadura inteligente a transmitir o teu nível de saúde. No entanto, isto resulta para todos os jogos? Não. Resulta para todo o tipo de informação? Não. O mesmo se passa com o reconhecimento de voz.”

Perguntei: Alguma vez implementas-te esta tecnologia enquanto *developer*?

Miguel hesita, dizendo que não, “A verdade é que nunca me lembrei de tal possibilidade e também teria medo de a implementar. Primeiro, porque acho que esta tecnologia ainda não está no ponto. Quero dizer, se esta falha muito então irá frustrar o jogador e não é isso que se quer. Mas isso pode ser a minha ideia, que pode estar ultrapassada. Não tenho noção do avanço que a tecnologia de reconhecimento de voz teve estes últimos anos, também por não ser uma coisa que use muito. Isso leva-me ao segundo ponto, o desconhecimento e falta de informação que existe a nível do seu potencial e de como o usar corretamente faz-me confusão.”

Miguel Fontoura ri-se, acabando por dizer, enquanto nos despedia-mos: “Sinceramente, após ter falado contigo, fiquei bastante interessado nesta tecnologia.

Eu trabalho com realidade virtual e o áudio é extremamente importante nesta área. Se uma pessoa se vir num mundo virtual, totalmente imerso nele, tornaria tudo muito mais natural se também ao falar pudesse ser ouvido e a sua voz afectar esse mundo. É realmente uma situação onde o uso da voz ficaria bem e fizesse sentido. Talvez seja algo interessante a ser estudado. Consigo ver possibilidades engraçadas ao juntar estas duas áreas.”

5.3. Apanhado das entrevistas

Segundo Miguel Fontoura, a tecnologia de reconhecimento de voz aplicada aos videojogos é realmente uma área na qual existe uma grande falta de conhecimento a ponto de causar desconforto aos *game developers* a ponto de estes ficarem abstraidos da possível implementação desta nos seus produtos.

De acordo com Mathieu Desautels, a tecnologia de reconhecimento de voz pode realmente ajudar os videojogos a tornarem-se mais atraentes ao jogador, maximizando a sua imersão, diversão e simplificando a visualização do jogo ao descartar botões quando implementada correctamente, o qual Miguel Fontoura sublinha ao dizer que, esta implementação, deve ser feita dentro de contexto e não apenas com o objectivo de descartar botões, este sendo apenas uma boa consequência.

A implementação desta tecnologia poderá ser mal feita se o *developer* não tiver cuidado com alguns pontos, como o modo como se utiliza a voz como interacção (momentos críticos, frases difíceis de decorar, uso da voz aproximado a situações reais, etc...), a possível falta de entendimento devido a possível sotaque, o tipo de jogo que é e o seu ritmo e a necessidade que o jogador tem de compreender o gameplay imediatamente e com pouco esforço.

6. TESTES

6. Testes

Tal como já referenciei, pretendo perceber como o reconhecimento de voz nos videojogos afecta a jogabilidade e o jogador, a nível de imersividade e divertimento.

É fácil obter dados concretos em relação à qualidade e eficiência do reconhecimento de voz, se este “entende” o comando do jogador ou não e se o executa correctamente. Quanto à “imersividade” e “divertimento”, é um pouco mais difícil avaliar, no entanto é possível ter a ideia geral através dos comentários e comportamento do jogador.

Assim, para avaliar tudo isto, os testes de usabilidade são essenciais. Mas como fazer estes testes e a quem?

Nas páginas a seguir explico tudo isso: Como decidi quem testar, como esses voluntários foram abordados e escolhidos, como planeei os testes e como os executei.

6.1. Personas

Visto que o que pretendo avaliar é a tecnologia de reconhecimento de voz nos jogos, os indivíduos que são mais indicados para testar as suas potencialidades e limitações serão baseadas nas seguintes condições:

- A experiência enquanto jogador. (Jogador Proficiente / Jogador Regular / Jogadora Casual)
- A fluência em Inglês. (Sendo que o nível mínimo é o de conseguir entender o suficiente para seguir instruções e compreender mensagens do jogo para que o consiga jogar.)
- O sotaque. (Ausência / Algum / Bastante Pronunciado)
- Ter entre 18 a 30 anos de idade.

A escolha destas condições deve-se aos diferentes aspectos que podem afectar a utilização do reconhecimento de voz, sendo que o hábito ou a falta deste pode afectar o modo como o jogador interage com a tecnologia, enquanto que a fluência em Inglês e sotaque tem a possibilidade de influenciar a capacidade de interpretação do reconhecimento de voz.

A escolha das idades, entre 18 e 30 anos, deve-se ao facto de ser esta a faixa etária que mais contem jogadores ou indivíduos não jogadores que estão habituados a este conceito. O facto da idade mínima ser 18 anos foi por motivos práticos de maioria de idade para que não existissem problemas.

Assim, tendo estes factores em conta, foram criadas as seguintes Personas, as quais ordenei em tabelas e que podem ser vistas asseguir, nas figuras

André	
<ul style="list-style-type: none"> - 18 anos - Português - Estudante, 12º ano - Vive com a avó 	<p>-Entende Instruções na língua inglesa, orais e escritas.</p> <p>- Comunica melhor em inglês por escrito do que oralmente por falta de uso.</p> <p>- Possui sotaque português que afecta um pouco a sua comunicação oral em inglês.</p>
<p>Aos 14 anos, recebeu o seu primeiro computador e tem, desde então, passado muito do seu tempo livre a jogar video-jogos de acção e táticos.</p>	<p>Nunca utilizou reconhecimento de voz nos seus video-jogos, falando apenas para comunicar com outros jogadores durante o jogo.</p>

Tabela 1 – Persona André

Sara	
<ul style="list-style-type: none"> - 30 anos - Inglesa - Restauradora num Museu - Vive com o marido e dois filhos. 	<p>-Inglês é a sua língua materna, logo é bastante fluente.</p> <p>- Possui sotaque britânico leve, o seu inglês sendo bastante perceptível.</p>
<p>Nunca ligou muito à área de jogos, no entanto, gosta de alguns jogos de simulação. Raramente joga e, desde que casou e teve filhos, que não tem tempo, preferindo jogar coisas pequenas como os jogos no seu Facebook.</p>	<p>Nunca utilizou reconhecimento de voz nos video-jogos, no entanto já conhece a tecnologia de outras áreas.</p>

Tabela 2 – Persona Sara

Valdir	
<ul style="list-style-type: none">- 26 anos- Russo- Programador- Vive com colega de trabalho.	<ul style="list-style-type: none">-Fluente em inglês, tanto a ler como a escrever e a falar.- Possui bastante sotaque russo, o que afecta imenso a sua comunicação oral em inglês.
<p><i>Desde pequeno que tem acesso a computadores e consolas, investindo bastante do seu tempo, ao longo da sua vida, a jogar os mais variados tipos de video-jogos, participando, até, em competições, sozinho ou com amigos.</i></p>	<p><i>Devido à sua experiência, já chegou, várias vezes, a experimentar video-jogos que utilizam reconhecimento de voz e necessitam do uso da fala para poderem ser jogados.</i></p>

Tabela 3 – Persona Valdir

6.2. Jogos escolhidos

“There came an Echo” da Iridium Studios [6]

Este é um videojogo de ficção científica de estratégia a tempo-real com uma narrativa complexa, na qual o jogador pode utilizar a voz para comandar as suas unidades pelo campo de guerra de modo a manter a sua vantagem táctica sobre o inimigo.

“In Verbis Virtus” da Indomitus Games [7]

Este é um videojogo de aventura jogado em primeira pessoa que mistura acção e quebra-cabeças num mundo de fantasia. Através da voz, o jogador recita encantamentos para fazer feitiços e ultrapassar os desafios que encontra.

Estes dois videojogos foram escolhidos devido ao seu fácil acesso e custo reduzido e por serem diferentes um do outro a nível de tema, *gameplay* e o fim para o qual ambos utilizam o reconhecimento de voz. Devido a serem jogos simples, não necessitam de muito tempo a serem jogados para testar o reconhecimento de voz pois o momento da utilização desta tecnologia ocorre logo desde início.

Tanto o *“In Verbis Virtus”* como o *“There came an Echo”* são títulos de jogos criados a partir de um projecto independente, ou seja, por pequenas equipas com pouco ou nenhum apoio financeiro de publicadoras e que frequentemente se focam em inovar a indústria de videojogos, normalmente não lhes sendo impostas muitas limitações a nível de como o jogo deve ser feito o que faz com que tenham mais liberdade em criar videojogos fora do normal. Este ponto foi outro motivo que me levou a escolhe-los.

6.3. Condições e Procedimentos dos testes

Para o desenvolvimento deste trabalho, é necessário um espaço onde se possa manter o silêncio e também recriar algum barulho e conversa de fundo para testar o reconhecimento de voz em ambas as situações, tendo utilizado o Laboratório de Usabilidade da Escola Superior de Educação de Coimbra do Instituto Politécnico de Coimbra. O portátil terá que suportar os jogos para que não haja interferência a nível do sistema, como lentidão, etc..., e terá que possuir microfone, para que se faça os testes com este e outro adicional com maior qualidade para uma comparação de resultados, tendo sido utilizado o microfone de uns auscultadores específicos para jogadores de videojogos. Durante os testes, é necessário uma câmara para gravar o *Think-Aloud* dos voluntários, as suas expressões e os movimentos que fazem com o rato/teclado e um *software* de recolha de imagem para a gravação do jogo.

Cada participante fez quatro testes de 20 a 23 minutos, participando numa sessão de uma hora e meia, na qual os participantes não estavam isolados uns dos outros, havendo um ambiente casual e sem pressão, de modo a que pudesse haver diálogo para a produção de ruído em alguns testes. Foi também pedido ao participante que estivesse a jogar para ir comentando o que estava a fazer e a pensar durante o jogo.

O primeiro teste foi com o videojogo “*In Verbis Virtus*” utilizando apenas o microfone do portátil, no qual o participante tinha que passar os seus desafios pronunciando encantamentos para produzir feitiços, clicando no rato para que o sistema do jogo começasse a detectar voz e largando o clique para produzir o feitiço. Assim, ao longo do jogo, foi medido o tempo que demorava ao participante pronunciar o encantamento, começando a contar desde o momento em que clicava até ao momento que largava o clique, anotando se o comando era executado pelo videojogo, produzindo o feitiço correcto, ou se falhava, não produzindo nada ou produzindo o feitiço errado. Enquanto o teste é feito, os participantes foram encorajados a dialogar de modo a criar algum ruído para testar o reconhecimento de voz. A ideia seria fazer outros testes, um em silêncio e outro com o microfone dos auscultadores, mas tal foi desnecessário, apesar de um pequeno teste para o

confirmar, pois o reconhecimento de voz deste jogo não era afectado pela qualidade do microfone nem pelo ruído.



Fig.0 - Participante a testar o videojogo *"In Verbis Virtus"* no laboratório de usabilidade, utilizando apenas o microfone do portátil.

O Segundo teste foi feito com o videojogo *"There came an Echo"* utilizando apenas o microfone do portátil. Neste jogo, o próprio jogador assume o papel de Sam, que, através de comandos por voz, ajuda, direciona e lidera outras personagens que estão em campo. Assim, neste jogo, foi medido o tempo que demorava ao participante pronunciar o comando, começando a contar desde o momento em que começava a falar até ao momento que o videojogo produzia um pequeno som, avisando que tinha reconhecido o comando e executando-o, tomando nota da quantidade de vezes que o participante se teve que repetir até ser reconhecido. Este teste foi feito em silêncio, pois foi óbvio desde início que o sistema tinha bastantes dificuldades a reconhecer a voz do jogador quando existia ruído, tornando-se quase impossível continuar o teste.

O Terceiro teste foi feito também com o videojogo *“There came an Echo”*, desta vez utilizando apenas o microfone dos auscultadores, o qual detectava apenas o participante, o ruído não afectando o resultado. Este teste foi executado exactamente como o segundo teste, tendo sido utilizado o mesmo método para medir o tempo que o participante demorava a pronunciar o comando e tendo sido também anotadas as repetições feitas até esse comando ter sido reconhecido.



Fig.1 - Participante a testar o videojogo *“There came an Echo”* no laboratório de usabilidade, utilizando apenas o microfone do portátil.

O Quarto e último teste, feito também com o videojogo *“There came an Echo”*, mas desta vez ignorando o sistema de reconhecimento de voz e utilizando apenas o rato para dar comandos. Neste teste foi medido o tempo que o participante demorava a dar o comando às personagens através do uso do rato, sendo medido apartir do momento que o jogador mexia o rato para escolher a personagem, encontrando o comando, clicando nele, no qual se ouve a voz de uma personagem do videojogo a dar a ordem, e terminando a contagem quando se ouvia o som produzido pelo videojogo, avisando que tinha reconhecido o comando e executando-o, não sendo necessário anotar mais valor nenhum pois este era sempre executado com sucesso e logo à primeira.

A necessidade de medir o *delay*, ou seja, o tempo que demorava ao sistema a executar o comando assim que este era percebido, foi desnecessário pois em ambos os jogos este era mínimo ou inexistente. Assim que se largava o rato, no caso do videojogo “*In Verbis Virtus*”, o comando era logo executado. No caso do “*There came an Echo*”, assim que reconhecia o comando de voz do jogador, o videojogo produzia imediatamente um som a anunciar que o comando tinha sido detectado e este era logo executado.

Após os quatro testes com o participante, era-lhe feito, sem a presença dos outros participantes, um pequeno inquérito e uma entrevista livre para fazer um apanhado da sua opinião, pensamento e sentimento acerca da utilização do reconhecimento de voz nos videojogos, querendo saber principalmente se o participante se tinha sentido frustrado ou divertido e imerso no videojogo e se tal foi afectado, ou não, pelo reconhecimento de voz, seja positivamente ou negativamente.

O inquérito tinha como base afirmações com as quais os participantes concordavam ou negavam através da escolha de “Sim” ou “Não”, não limitando os participantes a poderem acrescentar comentários que achassem relevantes. As afirmações focavam-se na maneira como os participantes se sentiram ao longo do jogo e como este o afectou, tendo estes que responder ao inquérito duas vezes, uma para cada jogo. As afirmações feitas foram as seguintes:

- Perdi a noção do tempo.
- Perdi a noção de onde estava (fora do jogo).
- Senti-me abstraído do mundo real.
- Senti-me imerso no jogo.
- Joguei sem ter que pensar como jogar.
- Foi natural para mim usar a voz para jogar.
- Usei voz para jogar sem ter que pensar no modo como proferir os comandos.

- Usar a voz ajudou a manter-me interessado no jogo.
- Usar a voz fez-me querer parar de jogar.
- Senti-me frustrado enquanto jogava.
- Senti-me calmo enquanto jogava.
- Senti-me entusiasmado enquanto jogava.
- Senti que queria jogar mais tempo.
- Fiquei interessado em jogar mais jogos que usem reconhecimento de voz para jogar.

6.4. Participantes

Através dos meios sociais e da internet, foi feita uma chamada a todos os que quisessem participar e ajudar nos testes, ficando apenas dez após alguns terem sido filtrados através de uma pequena entrevista, utilizando as personas criadas como guia e a disponibilidade de cada um como factores de escolha.

Os voluntários que participaram nesta experiência foram os seguintes:

Primeiro Participante: Sexo feminino, 25 anos, portuguesa, inglês como linguagem nativa, sem sotaque e jogadora proficiente. Não conhecia nenhum dos videojogos e já tinha utilizado a tecnologia de reconhecimento de voz antes.

Segundo Participante: Sexo masculino, 23 anos, português, fluente a inglês, com algum sotaque e jogador proficiente. Não conhecia nenhum dos videojogos e já tinha utilizado a tecnologia de reconhecimento de voz antes.

Terceiro Participante: Sexo feminino, 23 anos, portuguesa poucos conhecimentos da linguagem inglesa, com sotaque bastante pronunciado e jogadora casual. Não conhecia nenhum dos videojogos e nunca usou a tecnologia de reconhecimento de voz.

Quarto Participante: Sexo masculino, 27 anos, sueco, fluente a inglês, com sotaque bastante pronunciado e jogador regular. Conhecia apenas o *“In Verbis Virtus”* e já tinha utilizado a tecnologia de reconhecimento de voz antes.

Quinto Participante: Sexo feminino, 30 anos, portuguesa, bons conhecimentos da linguagem inglesa, sotaque bastante pronunciado, tendo uma limitação na fala devido a um problema no maxilar, jogadora casual. Não conhecia nenhum dos videojogos e nunca usou a tecnologia de reconhecimento de voz antes.

Sexto Participante: Sexo masculino, 28 anos, português, fluente a inglês, sem sotaque, jogador regular. Não conhecia nenhum dos videojogos e já tinha utilizado tecnologia de reconhecimento de voz antes.

Sétimo Participante: Sexo masculino, 30 anos, português, bons conhecimentos da linguagem inglesa, algum sotaque, jogador casual. Não conhecia nenhum dos videojogos e nunca usou a tecnologia de reconhecimento de voz antes.

Oitavo Participante: Sexo masculino, 18 anos, português, fluente a inglês, bastante sotaque, jogador proficiente. Não conhecia nenhum dos videojogos, nunca tinha utilizado tecnologia de reconhecimento de voz antes.

Nono Participante: Sexo feminino, 20 anos, portuguesa, bons conhecimentos da linguagem inglesa, algum sotaque, jogadora regular. Conhecia o “In Verbis Virtus” de nome, nunca utilizou tecnologia de reconhecimento de voz antes.

Décimo Participante: Sexo masculino, 27 anos, holandês, fluente a inglês, algum sotaque, jogador proficiente. Não conhecia nenhum dos videojogos, nunca tinha utilizado tecnologia de reconhecimento de voz antes.

6.5. Hipóteses

Baseado nas informações que recolhi anteriormente e que foram apresentadas acima, formulei as seguintes hipóteses acerca do uso da tecnologia de reconhecimento de voz nos videojogos, esperando encontra-las durante e após os testes feitos com os participantes no laboratório de usabilidade.

Hipótese 1: Os jogadores habituados a jogar videojogos com o teclado e o rato irão utiliza-los com mais rapidez do que ao utilizarem o reconhecimento de voz.

Hipótese 2: O uso de linguagem fictícia faz com que o sistema de reconhecimento de voz tenha maior facilidade em compreender o jogador.

Hipótese 3: O reconhecimento de voz trás maior facilidade na execução de comandos em simultâneo.

Hipótese 4: Existirá uma curva de aprendizagem para com o sistema de reconhecimento de voz, sendo esperado que haja uma maior dificuldade inicial em funcionar com este.

6.6. Resultados

Foi verificado, no primeiro teste, que o reconhecimento de voz afectava o videojogo *“In Verbis Virtus”* de uma maneira positiva.

Foi observado que os participantes se divertiam, comentando positivamente o jogo e ficando surpreendidos com este. Durante os testes, os jogadores entraram no espírito de quererem testar a capacidade da tecnologia implementada no videojogo, experimentando dar comandos, ou seja, pronunciar os encantamentos de linguagem estranha e fictícia que tinham à sua disposição com diferentes tipos de voz e volume, tentando até enganar o sistema de reconhecimento de voz ao pronunciarem palavras parecidas com os encantamentos para ver se o sistema as executava de qualquer modo. Estas experimentações revelaram algo surpreendente, pois o sistema não deixou de entender o que o jogador dizia, executando sempre o feitiço certo e não executando quando as palavras eram erradas. A única excepção foi o quinto participante, que devido a um problema no maxilar, não conseguia pronunciar certos sons e acabava, por vezes, por não conseguir ser entendida pelo sistema de todo, tendo que ter mais atenção e esforço com a sua dicção.

Como podemos ver na Fig.2, os valores foram bastante consistentes entre os participantes e, apesar das suas diferenças a nível de sotaque e das experiências de voz (últimos quatro valores), podemos ver que o sistema de reconhecimento de voz acabou por conseguir executar acertadamente o comando dado na maioria das vezes, não tendo compreendido o quinto participante apenas três vezes devido à sua limitação. Em falta, os valores das experiências que os participantes fizeram a dizer palavras parecidas com os encantamentos que deveriam pronunciar, estes não tendo sido executados pelo sistema do videojogo, mostrando que este reconhecia bem as palavras, não deixando o jogador fazer batota.

In Verbis Virtus		In Verbis Virtus		In Verbis Virtus		In Verbis Virtus		In Verbis Virtus	
Participante 1		Participante 2		Participante 3		Participante 4		Participante 5	
Tempo	(E)xecutado ou (N)ão	Tempo	(E)xecutado ou (N)ão	Tempo	(E)xecutado ou (N)ão	Tempo	(E)xecutado ou (N)ão	Tempo	(E)xecutado ou (N)ão
2.48s	E	2.16s	E	2.89s	E	2.25s	E	2.09s	N
2.80s	E	2.81s	E	2.00s	E	2.03s	E	2.00s	N
2.27s	E	2.00s	E	2.47s	E	2.26s	E	2.02s	N
2.30s	E	2.56s	E	2.31s	E	2.15s	E	2.43s	E
2.51s	E	2.05s	E	2.44s	E	2.65s	E	2.00s	N
2.56s	E	2.54s	E	2.20s	E	2.22s	E	2.42s	E
2.12s	E	2.12s	E	2.57s	E	2.36s	E	2.34s	E
2.03s	E	2.54s	E	2.26s	E	2.00s	E	2.40s	E
2.16s	E	2.50s	E	2.32s	E	2.26s	E	2.41s	E
2.13s	E	2.00s	E	2.12s	E	1.98s	E	2.10s	E
In Verbis Virtus		In Verbis Virtus		In Verbis Virtus		In Verbis Virtus		In Verbis Virtus	
Participante 6		Participante 7		Participante 8		Participante 9		Participante 10	
Tempo	(E)xecutado ou (N)ão	Tempo	(E)xecutado ou (N)ão	Tempo	(E)xecutado ou (N)ão	Tempo	(E)xecutado ou (N)ão	Tempo	(E)xecutado ou (N)ão
2.33s	E	2.45s	E	2.42s	E	2.20s	E	2.83s	E
2.15s	E	2.25s	E	2.78s	E	2.09s	E	2.12s	E
2.37s	E	2.32s	E	2.27s	E	2.23s	E	2.40s	E
2.37s	E	2.43s	E	2.53s	E	2.26s	E	2.00s	E
2.31s	E	2.59s	E	2.45s	E	2.52s	E	2.24s	E
2.45s	E	2.23s	E	2.29s	E	2.39s	E	2.23	E
2.20s	E	2.49s	E	2.17s	E	2.33s	E	2.25s	E
2.22s	E	2.11s	E	2.31s	E	2.27s	E	2.27s	E
2.27s	E	2.34s	E	2.21s	E	2.42s	E	2.20s	E
2.23s	E	2.09s	E	2.10s	E	2.20s	E	2.14s	E

Fig.2 - (Primeiro Teste, “In Verbis Virtus”, microfone do portátil): Tempo que cada participante demorou a pronunciar a palavra e se o comando foi compreendido/executado ou não pelo videojogo.

No segundo, terceiro e quarto teste, feitos com o videojogo “There came an Echo”, foi verificado que o reconhecimento de voz afectava o videojogo de uma forma negativa tendo, no entanto, alguns pontos positivos.

Durante estes testes, os participantes largaram o espírito de experimentação, querendo apenas tentar que o reconhecimento de voz funcionasse, devido às frustrações que o sistema criava ao jogador.

Foi verificado imediatamente a necessidade de silêncio para que o sistema de reconhecimento de voz funcionasse.

Como podemos ver na Fig.3, todos os participantes tiveram uma dificuldade inicial em lidar com o sistema, como era esperado, aprendendo e adaptando-se ao longo do jogo, conseguindo uma maior taxa de sucesso nas últimas experiências.

É de sublinhar a dificuldade que se volta a ver no quinto participante, chegando até a desistir de dar o comando e, fazendo “batota”, utilizou o rato para passar à frente e continuar o teste.

Curiosamente, o Participante 2 teve dificuldade em fazer o videojogo executar os comandos, apesar de ser um dos que tinha a melhor pronúncia inglesa, havendo a possibilidade do sistema ter maior dificuldade em detectar a sua frequência e tom de voz. Neste teste, os participantes ficaram bastante frustrados, queixando-se muito do jogo e ficando impacientes, focando-se em conseguir ser entendidos pelo reconhecimento de voz em vez de simplesmente jogarem o jogo, tornando o sistema de voz em algo nada natural nem familiar. O único a quem o jogo correu bastante bem e sem queixas, suavemente e sem problemas, foi ao sexto participante, o qual tinha a voz mais grave, calma e clara de todos os participantes.

There came an Echo		There came an Echo		There came an Echo		There came an Echo		There came an Echo	
Participante 1		Participante 2		Participante 3		Participante 4		Participante 5	
Tempo	Repetição até sucesso	Tempo	Repetição até sucesso	Tempo	Repetição até sucesso	Tempo	Repetição até sucesso	Tempo	Repetição até sucesso
8.10s	3x	18.78s	5x	21.23s	7x	13.43s	3x	25.12s	10x (Desistiu)
8.01s	2x	20.12s	6x	19.78s	6x	8.52s	2x	20.00s	6x
3.86s	1x	15.12s	4x	3.15s	2x	8.45s	2x	17.24s	5x
4.09s	1x	3.75s	2x	4.75s	3x	4.41s	1x	15.10s	8x (Desistiu)
3.82s	1x	3.25s	2x	1.13s	1x	4.07s	2x	8.37s	4x
3.72s	1x	5.79s	3x	1.28s	1x	3.91s	1x	8.12s	4x
7.67s	2x	1.10s	1x	3.89s	2x	4.73s	2x	4.87s	2x
6.96s	2x	21.23s	6x	1.10s	1x	5.24s	2x	15.30s	5x (Desistiu)
6.13s	2x	14.10s	4x	1.12s	1x	3.70s	1x	9.15s	3x
3.97s	1x	3.68s	2x	4.25s	2x	3.41s	1x	4.60s	2x

There came an Echo		There came an Echo		There came an Echo		There came an Echo		There came an Echo	
Participante 6		Participante 7		Participante 8		Participante 9		Participante 10	
Tempo	Repetição até sucesso	Tempo	Repetição até sucesso	Tempo	Repetição até sucesso	Tempo	Repetição até sucesso	Tempo	Repetição até sucesso
8.31s	3x	9.76s	3x	13.90s	4x	14.12s	4x	8.14s	3x
8.40s	2x	9.23s	2x	12.09s	3x	15.34s	4x	8.55s	2x
4.02s	1x	3.68s	2x	4.46s	3x	3.80s	2x	6.80s	3x
3.82s	1x	3.25s	1x	3.78s	2x	3.70s	1x	4.00s	1x
1.56s	1x	4.99s	3x	3.76s	2x	4.20s	2x	1.56s	1x
1.30s	1x	2.67s	2x	4.13s	3x	1.99s	1x	2.02s	1x
1.34s	1x	2.00s	1x	2.17s	1x	3.90s	2x	4.67s	2x
2.21s	2x	3.30s	2x	2.12s	1x	5.50s	2x	2.40s	2x
1.17s	1x	4.85s	2x	3.15s	1x	3.50s	1x	1.20s	1x
1.48s	1x	3.78s	2x	2.20s	1x	3.49s	2x	2.00s	1x

Fig.3 - (Segundo Teste, "There came an Echo", microfone do portátil): Tempo que cada participante demorou a pronunciar a palavra e a quantidade de vezes que a teve de repetir até ser compreendida/executada pelo videojogo.

No terceiro teste, passámos à utilização do microfone dos auscultadores. Neste teste, os participantes relaxaram um pouco mais e conseguiram utilizar a tecnologia de reconhecimento de voz de uma maneira mais natural ao repararem que o sistema conseguia reconhecer os seus comandos com maior facilidade do que anteriormente. Podemos então assumir que é necessário a utilização de auscultadores

com microfone para que o sistema do videojogo funcione e o possamos aproveitar ao máximo, algo que o próprio jogo sugere ao jogador no início do videojogo.

There came an Echo (Auscultadores)		There came an Echo (Auscultadores)		There came an Echo (Auscultadores)		There came an Echo (Auscultadores)		There came an Echo (Auscultadores)	
Participante 1		Participante 2		Participante 3		Participante 4		Participante 5	
Tempo	Repetição até sucesso	Tempo	Repetição até sucesso	Tempo	Repetição até sucesso	Tempo	Repetição até sucesso	Tempo	Repetição até sucesso
4.02s	1x	15.14s	4x	5.07s	2x	8.67s	2x	25.80s	10x (Desistiu)
3.78s	1x	14.98s	4x	5.70s	3x	8.35s	2x	23.87s	8x (Desistiu)
3.87s	1x	5.66s	3x	4.67s	2x	4.40s	1x	5.43s	3x
7.00s	2x	3.84s	2x	4.38s	2x	3.43s	1x	12.23s	5x (Desistiu)
6.96s	2x	3.82s	1x	2.20s	1x	3.79s	1x	4.77s	2x
2.70s	1x	5.50s	3x	3.76s	2x	2.30s	1x	6.08s	4x
2.66s	1x	1.04s	1x	1.29s	1x	2.50s	2x	3.19s	3x
1.69s	1x	4.05s	2x	1.10s	1x	2.03s	2x	3.67s	2x
3.98s	2x	3.79s	2x	1.22s	1x	1.67s	1x	4.32s	3x
1.96s	1x	4.68s	3x	1.12s	1x	1.82s	1x	5.44s	4x

There came an Echo (Auscultadores)		There came an Echo (Auscultadores)		There came an Echo (Auscultadores)		There came an Echo (Auscultadores)		There came an Echo (Auscultadores)	
Participante 6		Participante 7		Participante 8		Participante 9		Participante 10	
Tempo	Repetição até sucesso	Tempo	Repetição até sucesso	Tempo	Repetição até sucesso	Tempo	Repetição até sucesso	Tempo	Repetição até sucesso
4.19s	1x	4.20s	1x	4.32s	1x	3.97s	1x	8.33s	3x
4.01s	1x	3.59s	1x	3.70s	1x	5.33s	3x	5.90s	2x
3.22s	1x	4.30s	2x	3.41s	1x	3.21s	1x	5.47s	2x
3.88s	1x	3.94s	1x	4.40s	2x	4.00s	2x	4.00s	1x
2.16s	1x	2.80s	1x	7.24s	3x	2.67s	1x	2.12s	1x
2.50s	1x	3.52s	2x	2.86s	1x	2.40s	1x	3.80s	2x
1.15s	1x	2.28s	1x	2.66s	2x	2.16s	1x	1.99s	1x
3.09s	2x	3.85s	2x	1.79s	1x	1.55s	1x	2.05s	1x
1.29s	1x	1.57s	1x	1.48s	1x	2.04s	1x	1.23s	1x
1.14s	1x	2.20s	2x	3.00s	2x	1.71s	1x	1.87s	1x

Fig.4 - (Terceiro Teste, “There came an Echo”, microfone dos auscultadores): Tempo que cada participante demorou a pronunciar a palavra e a quantidade de vezes que a teve de repetir até ser compreendida/executada pelo videojogo.

Olhando para a Fig.4 e comparando-a com as tabelas na Fig.3, podemos observar a diminuição na necessidade de repetição para o sistema reconhecer e executar o comando, existindo neste terceiro teste ausência da dificuldade inicial existente no teste anterior, havendo a possibilidade desta situação não ter acontecido devido ao bom microfone ou porque os participantes já estavam habituados ao reconhecimento de voz. É de sublinhar que também neste teste o Participante 2 teve alguma dificuldade em ser reconhecido, apesar desta ter sido diminuída pelo uso de auscultadores. Quanto ao quinto participante, a sua dificuldade manteve-se e a mudança de valores do segundo para o terceiro teste foi mínima.

Como o videojogo “There came an Echo” possibilitava a utilização de rato para a execução de comandos, houve a possibilidade de também testar este sistema e

compará-lo à sua tecnologia de voz. Assim, no quarto teste, os participantes jogaram apenas com o rato, voltando à conversa animada. Durante este teste, os participantes não tiveram qualquer dificuldade nem frustração, tendo havido apenas um comentário negativo a partir de dois dos participantes dizendo que, passo a citar, “O jogo é bem interessante, mas assim não tem tanta piada, passa a ser mais um entre os outros!”. Quanto ao quinto participante, afirmou determinadamente que o jogo era melhor sem a utilização de voz.

There came an Echo (Rato)	There came an Echo (Rato)	There came an Echo (Rato)	There came an Echo (Rato)	There came an Echo (Rato)
<i>Participante 1</i>	<i>Participante 2</i>	<i>Participante 3</i>	<i>Participante 4</i>	<i>Participante 5</i>
Tempo	Tempo	Tempo	Tempo	Tempo
8.07s	7.40s	8.02s	7.04s	7.45s
6.61s	6.76s	7.10s	6.83s	7.00s
4.29s	4.89s	6.60s	4.42s	4.90s
5.83s	5.10s	5.57s	4.06s	5.64s
5.14s	4.52s	5.36s	5.60s	4.93s
4.65s	4.36s	5.12s	4.62s	4.39s
5.30s	4.87s	4.98s	5.48s	4.94s
4.36s	4.24s	5.01s	4.45s	4.38s
4.85s	5.09s	4.55s	4.77s	5.30s
4.60s	4.78s	4.36s	4.54s	4.60s

There came an Echo (Rato)	There came an Echo (Rato)	There came an Echo (Rato)	There came an Echo (Rato)	There came an Echo (Rato)
<i>Participante 6</i>	<i>Participante 7</i>	<i>Participante 8</i>	<i>Participante 9</i>	<i>Participante 10</i>
Tempo	Tempo	Tempo	Tempo	Tempo
7.40s	8.00s	7.52s	7.42s	8.02s
6.59s	6.60s	6.75s	6.76s	7.04s
4.63s	4.51s	4.60s	5.02s	4.80s
5.42s	5.77s	5.12s	5.55s	5.63s
4.89s	5.43s	5.82s	5.08s	4.99s
4.67s	4.67s	4.62s	4.81s	4.57s
5.00s	5.23s	5.18s	4.99s	5.00s
4.27s	4.42s	4.44s	4.18s	4.20s
4.31s	4.79s	4.57s	4.16s	2.22s
4.63s	4.57s	4.56s	4.71s	4.59s

Fig.5 - (Quarto teste, “There came an Echo”, rato): Tempo que cada participante demorou a utilizar o rato para executar o comando no videojogo.

Tal como mostram as tabelas na Fig.5, o tempo torna-se mais consistente. Apesar de este ser um pouco mais elevado que nos outros testes, devido ao percurso

que o rato tem que fazer no ecrã e o tempo que demora ao jogador a encontrar onde quer clicar, o comando é sempre executado com sucesso e o jogador não tem noção que está realmente a demorar um pouco mais, o gameplay do videojogo sendo fluído.

Apesar destes dados darem bastante informação, a par com os comentários dos participantes, não demonstram o sentimento do participante enquanto jogava cada um dos jogos. Assim, com a ajuda de um questionário, no qual afirmações eram feitas e com as quais o jogador concordava ou negava, foi possível entender o tipo de emoção e sentimento de maior força entre os participantes para com cada um dos videojogos, a nível do reconhecimento de voz.

	There came an Echo		In Verbis Virtus	
	<i>Sim</i>	<i>Não</i>	<i>Sim</i>	<i>Não</i>
Perdi a noção do tempo.	30%	70%	90%	10%
Perdi a noção de onde estava (fora do jogo).	30%	50%	40%	60%
Senti-me abstraído do mundo real.	50%	50%	80%	20%
Senti-me imerso no jogo.	10%	90%	90%	10%
Joguei sem ter que pensar como jogar.	30%	70%	50%	50%
Foi natural para mim usar a voz para jogar.	10%	90%	80%	20%
Usei voz para jogar sem ter que pensar no modo como proferir os comandos.	30%	70%	80%	20%
Usar a voz ajudou a manter-me interessado no jogo.	40%	60%	80%	20%
Usar a voz fez-me querer parar de jogar.	70%	30%	0%	100%
Senti-me frustrado enquanto jogava.	80%	20%	20%	80%
Senti-me calmo enquanto jogava.	40%	60%	80%	20%
Senti-me entusiasmado enquanto jogava.	60%	40%	100%	0%
Senti que queria jogar mais tempo.	40%	60%	60%	40%
Fiquei interessado em jogar mais jogos que usem reconhecimento de voz para jogar.	40%	60%	80%	20%

Fig.6 – (Respostas dadas pelos 10 participantes, concordando ou negando as afirmações para cada um dos videojogos)

É possível ver, olhando para a Fig.6, que ambos os jogos afectaram os participantes de maneiras diferentes e, apesar as respostas serem diferentes de participante para participante e de jogo para jogo, pode-se ver a existência de uma tendência, a qual é possível analisar através de uma média, demonstrada na Fig.7.

	Média de ambos os jogos	
	Sim	Não
Perdi a noção do tempo.	60%	40%
Perdi a noção de onde estava (fora do jogo).	35%	65%
Senti-me abstraído do mundo real.	65%	35%
Senti-me imerso no jogo.	50%	50%
Joguei sem ter que pensar como jogar.	40%	60%
Foi natural para mim usar a voz para jogar.	45%	55%
Usei voz para jogar sem ter que pensar no modo como proferir os comandos.	55%	45%
Usar a voz ajudou a manter-me interessado no jogo.	60%	40%
Usar a voz fez-me querer parar de jogar.	35%	65%
Senti-me frustrado enquanto jogava.	50%	50%
Senti-me calmo enquanto jogava.	60%	40%
Senti-me entusiasmado enquanto jogava.	80%	20%
Senti que queria jogar mais tempo.	50%	50%
Fiquei interessado em jogar mais jogos que usem reconhecimento de voz para jogar.	60%	40%

Fig.7 – (Média das respostas dadas aos dois videojogos, pelos 10 participantes, concordando ou negando as afirmações)

Podemos ver, na Fig.7, que houve uma grande maioria sentiu-se entusiasmado com os videojogos e, apesar de haver apenas metade a dizer que se sentia imerso neles, a verdade é que a percentagem maior acabou por se sentir abstraído do mundo real e perdeu a noção do tempo enquanto jogava.

A utilização da voz nos videojogos acabou por, de acordo com as respostas dos participantes, tornar os videojogos mais interessantes para muitos deles, apesar de haver vários participantes que não acharam o uso desta tecnologia como modo para jogar muito natural, justificando-se com o facto de não estarem habituados, “Não é uma coisa comum que se encontre nos jogos, acabo por estar sempre consciente da necessidade de falar. Ou seja, tenho que me lembrar que para interagir com o jogo é assim e não com o teclado e rato como costume.” (Participante 4)

No entanto, voltando a olhar para a Fig.6, podemos ver que o reconhecimento de voz teve um impacto bastante positivo no jogo “*In Verbis Virtus*”, enquanto no “*There came an Echo*” este acabou por se tornar frustrante, havendo participantes que acharam que o reconhecimento de voz os atrapalhou e fez com que quisessem parar de jogar.

6.6.1. Resumo dos resultados

De acordo com as hipóteses previstas, situações que foram assumidas serem vistas durante os testes, foram verificadas como correctas a segunda, terceira e quarta hipótese. Na primeira hipótese assumi que os jogadores habituados ao rato e ao teclado usariam estes periféricos mais rapidamente do que o uso da tecnologia de reconhecimento de voz, no entanto nem sempre tal situação ocorreu, apesar de, no uso das teclas, ter havido um tempo de utilização bastante estável, o que torna o seu uso confortável. Mas quer isto dizer que é mais prático? De acordo com Mathieu Desautels, com os participantes e pesquisa feita, mesmo que o reconhecimento de voz torne o gameplay mais vagaroso, se o sistema estiver a funcionar correctamente, o jogador não se vai importar de sacrificar tal valor (o tempo) para poder ter um modo de interacção mais natural e mais divertido.

Entre os videojogos “*In Verbis Virtus*” e “*There Came an Echo*”, pode-se dizer que a implementação da tecnologia de reconhecimento de voz como modo de interacção no videojogo está melhor “*In Verbis Virtus*”, pois este não só teve maior sucesso a nível de compreensão e execução de comandos mas também porque foi o videojogo no qual os participantes se divertiram mais e onde estes se sentiram entusiasmados e mais imersos, valor ao qual devemos dar a maior importância.

A razão de existir tão grande diferença de respostas, a nível do questionário, entre os dois videojogos tem a ver, não só com o modo como o reconhecimento de voz foi implementado e sua qualidade mas também com o tipo de jogo que era.

De acordo com alguns participantes, o videojogo “*There came an Echo*”, no qual tinham que dar comandos a uma equipa durante confrontos e batalhas com o inimigo, tornava-se bastante stressante, pois queriam dar as ordens correctas e serem entendidos dentro do tempo, às vezes a voz sendo afectada pelo entusiasmo, *stress* ou frustração que o jogador sentiu, elevando-a a ponto do sistema deixar de compreender as palavras do participante. O Participante 10 diz que “Fiquei bastante nervoso ao ver uma das personagens quase a morrer, queria tira-la dali! E eu a ve-la a levar com tiros e a ignorar as minhas ordens para voltar para trás para poder abrigar-se! Tive que pausar para dar as ordens, funcionou, mas penso que isso acabou por

tirar a imersividade do jogo.”. Os participantes 1, 6 e 7, comentaram também que, neste jogo, “A maneira como tinha-mos de dizer as frases para dar uma ordem era tão limitante! Havia apenas uma maneira correcta e por vezes esquecia-me do tipo de palavras ou gramática que tinha que usar para que o sistema acabasse por me entender!”.

Apesar de existirem vários participantes a dizer que o jogo *“In Verbis Virtus”* não era o seu estilo de jogo, afirmaram que este tinha um gameplay mais atraente e tudo parecia mais natural.

6.7. Heurísticas de Nielsen aplicadas aos jogos testados

As heurísticas de Nielsen são um guia que, apesar de estar definido para softwares e não para video-jogos, continua a dar-nos alguma informação importante, principalmente no que conta à interface e não ao gameplay e mecânicas de jogo.

Assim, mostrando-as por pontos, tento aplica-las, não ao jogo em geral mas em contexto com o uso da voz enquanto modo de controlo do jogo.

1. Visibilidade do Status do Sistema;

Em ambos os jogos, o utilizador consegue facilmente perceber se o comando oral dado foi percebido pelo sistema, podendo ver a acção a acontecer logo de imediato.

Caso o sistema não entenda o comando, no jogo “In Verbis Virtus”, existem efeitos visuais e sons que alertam para a situação. No entanto, no jogo “There came an Echo”, este não dá sinal de falha e mantém-se em silêncio, algo que traz alguma confusão ao jogador e, consequentemente, alguma frustração. Por outro lado, este tem um pequeno histórico no canto superior direito que indica todos os comandos executados que, para além de nos confirmar o sucesso da acção, mantém-nos a par da situação em que estamos.

2. Relacionamento entre a interface do sistema e o mundo real;

Em ambos os jogos, a comunicação a ser utilizada pelo utilizador para que o sistema o entenda não é familiar, o que vai contra esta heurística, no entanto, esta quebra é feita com intenção e é contextualizada pelo jogo, fazendo o jogador sentir que está realmente na posição e situação que a sua personagem do jogo está.

No jogo “In Verbis Virtus”, um jogo de fantasia, os comandos são encantamentos numa língua fictícia, palavras pequenas, que ficam no ouvido e que não tem demasiada variedade para que o jogador se lembre delas facilmente.

No jogo “There Came an Echo”, apesar da linguagem utilizada ser o Inglês, a maneira como é utilizada, ou seja, o facto desta ser mais formal e militar, algo que não é utilizado no dia-a-dia, torna-a menos familiar, mas à qual o jogador é introduzido lentamente para se habituar.

3. Liberdade e controlo do utilizador;

Em ambos os jogos, o jogador tem a liberdade de a qualquer momento poder dar um comando e de o poder desfazer ou contrariar, apesar de o jogo “There Came an Echo”, ter o controlo um pouco condicionado com a situação, a qual está em contexto com os eventos do jogo.

4. Consistência;

Ao longo de ambos jogos, a forma de dar comandos, fazer acções, resolver puzzles, entre outras coisas, mantêm um padrão, havendo sempre consistência. Os comandos por voz são dados sempre do mesmo modo e o que cada comando oral faz o sistema executar a acção apropriada, esperada pelo jogador.

5. Prevenção de erros;

A forma que estes jogos têm de prevenir erros, neste caso, de executar um comando quando não era suposto, é de limitar quando é que o sistema está a “ouvir” o utilizador.

No jogo “In Verbis Virtus”, o sistema apenas ouve o jogador quando este clica no rato e sabe que o utilizador terminou de proferir o comando quando este retira o

dedo do botão do rato, não havendo medo, por exemplo, do jogador poder estar a falar com alguém e o sistema interpretar alguma palavra como um encantamento acidentalmente.

No jogo “There Came an Echo”, o sistema apenas espera ouvir o utilizador quando o jogo avisa e dá sinal para o jogador falar ou durante batalhas, a qual o utilizador pode pausar facilmente.

6. Reconhecimento ao invés de lembrança;

Ambos os jogos, por usarem um tipo de comunicação não familiar, como já foi referido antes, vão um pouco contra esta heurística.

No jogo “In Verbis Virtus”, a memorização das seis palavras fictícias faz parte da imersão e desafio do jogo.

No jogo “There Came an Echo”, a utilização certa de como dar ordens não é fácil de lembrar, também fazendo parte da imersão do jogo, no entanto, a necessidade de memorização não faz qualquer parte do desafio, e sim, o modo como se utilizam essas ordens enquanto tática no jogo, fazendo esta necessidade uma pequena frustração para o jogador.

7. Flexibilidade e eficiência de uso;

Pode-se dizer que o jogo “In Verbis Virtus” não é nada flexível. Não existe qualquer tipo de customização. O utilizador é obrigado a utilizar os controlos e os comandos dados para fazer as acções e encantamentos, no entanto, essa falta de customização está em contexto com o jogo, onde o jogador é suposto saber aqueles comandos específicos, os quais são bastante eficientes e fáceis de usar.

O jogo “There came an Echo” é bastante flexível, possibilitando a customização de qualquer controlo ou comando oral, colmatando a possível falha de o jogador não se dar bem com o tipo de comunicação escolhida por defeito. É também de referir

que o jogador tem a capacidade de chegar a atalhos, através do uso do rato, para executar comandos em vez de utilizar o reconhecimento de voz como modo de controlo do jogo. Tudo isto mexe com a eficiência dos controlos/comandos e a execução destes pelo sistema.

8. Estética e design minimalista;

Em ambos os jogos, a utilização do reconhecimento de voz torna a interface simples e retira lixo visual, ou seja, informação desnecessária para o momento, a qual pode ser chamada a qualquer altura pelo jogador.

É de diferenciar o jogo “In Verbis Virtus”, que consegue o melhor design minimalista entre estes dois jogos, tendo apenas a necessidade de mostrar a barra de vida. Coisas como o diário, ajudas, etc., podem ser facilmente acedidas através de atalhos ou do menu principal, o qual é chamado pelo típico botão “Esc” (Escape).

9. Ajudar os utilizadores a reconhecer/diagnosticar/recuperar-se de erros;

O jogo “In Verbis Virtus” ajuda o utilizador a reconhecer que o comando oral falhou através de sinais visuais e de som. Apesar de não existir propriamente um diagnóstico, o jogador sabe que, se não existir sinais de falha após proferir o encantamento, é porque não clicou no rato para possibilitar o sistema de o “ouvir”. Se existir sinais de falha ou a acção executada for a errada é porque ou proferiu demasiado baixo ou erradamente. O jogador consegue recuperar-se deste erro ao tentar novamente, mais alto ou verificando o encantamento, pronuncia e o que faz, na documentação dada para o fazer da forma correcta.

O jogo “There Came an Echo” não ajuda a reconhecer os erros e muito menos a diagnosticar porque é que o comando falhou, podendo existir várias causas: Volume de voz, pronuncia, ordem das palavras, ordem dos comandos, etc... No entanto, o jogador consegue recuperar-se do erro ao utilizar os atalhos através do uso do rato.

10. Ajuda e documentação.

O jogo “In Verbis Virtus” têm ajuda de fácil acesso, o qual é introduzida ao jogador no início. Este tem um diário, que pode ser consultado quando o jogador quiser, para ler e ouvir a pronuncia dos encantamentos caso se esqueça destes e para se informar o que cada um faz e outros extras.

O jogo “There Came an Echo” tem documentação que ajuda o utilizador, no entanto, a ajuda que têm de mais fácil acesso é a capacidade de usar o rato para chamar atalhos, ou seja, comandos por escrito e, em vez de os utilizar para executar o comando, ler e saber que comandos pode proferir oralmente naquele momento e àquela unidade/personagem para que o sistema reconheça e execute a acção escolhida.

7. CONCLUSÕES

7. Conclusões

Podemos dizer, então, que a qualidade do sistema de reconhecimento de voz afecta o jogador e o videojogo. Mas para além disso, também a situação onde esta é utilizada e como o jogador a tem que utilizar pode arruinar o videojogo ou torna-lo um sucesso. Verificamos que esta tecnologia pode ajudar bastante na imersividade de um videojogo e aumentar o divertimento e confirmamos que este pode realmente ajudar numa interface mais limpa, deixando de existir informação visual a mais que desvie a atenção do jogador do próprio videojogo, o que ajuda também na imersão deste, no entanto, deve ser utilizado em contexto e não apenas com o objectivo de ter a tecnologia apenas para ter nem para retirar botões que fazem mais sentido enquanto botões.

Um dos motivos para o reconhecimento de voz não ser mais utilizado na industria de jogos é, de acordo com a informação que recolhi ao longo deste trabalho, a falta de informação que existe á volta desta tecnologia e, assim, a facilidade de cair em erro e de não a implementar correctamente, podendo arruinar o video-jogo desenvolvido.

7.1 Guia para a implementação do reconhecimento de voz nos videojogos

Através da informação recolhida ao longo deste relatório, podemos então sugerir alguns pontos para uma melhor implementação da tecnologia de reconhecimento de voz nos videojogos:

- Ter em conta se o jogo realmente precisa desta tecnologia, pois este pode afectar o videojogo negativamente caso não seja necessária. Ganha algum extra? Soluciona alguma coisa? Está a facilitar algo? Torna um comando/acção mais natural?
- Utilização do tipo de linguagem:
 - Caso seja linguagem fictícia:
 - Deve soar bem e ficar no ouvido, ou seja, ser memorizável.
 - Não devem ser utilizados sons que possam ser difíceis de pronunciar para pessoas de certos países, por exemplo o som “nh” e “lh” existentes no português.
 - Caso seja linguagem real:
 - As palavras que são exigidas ao jogador devem ser familiares.
 - O uso de respostas simples, por exemplo o “sim” e o “não”, tornam-se mais fáceis de usar para o jogador e mais fáceis de compreender pelo sistema.
 - Quando é exigido ao jogador a formulação de frases, estas devem ser bem pensadas a nível de estrutura para que sejam naturais e familiares e sejam ditas de forma correcta.
 - O sistema deve ter uma maior flexibilidade de compreensão na forma como as frases poderam ser construídas de maneira diferente entre os jogadores.
- Existir espaço de tempo para o jogador aprender a lidar com o sistema, habituar-se a este e saber o que é esperado de si.

- Ajudar o jogador, através de pistas, de como e o que dizer e mostrar-lhe as possibilidades que tem devagar, não afogando o jogador em informação.
- Ter o cuidado de arranjar forma de não existir delay entre o comando e sua execução/ momento em que o jogador espera que este seja executado.
- Tentar com que a utilização dos comandos vocais não seja demasiado repetitiva.
- Não deve ser utilizado em todas as situações nem em todos os tipos de videojogos:
 - Os videojogos que ganham mais com esta tecnologia são, geralmente, os jogos de ritmo lento.
 - Deve-se utilizar os comandos por voz em situações semelhantes a situações reais, para uma maior naturalidade do seu uso, por exemplo, para dar ordens ou recitar encantamentos.
 - Não ser utilizada para tudo mas sim para acções chave.
 - Deve ser usada caso haja necessidade de acções em simultâneo ou existência de inúmeros comandos/acções, combinadas ou não.
- Deve ser utilizada caso se queira chegar a um público diferente, como por exemplo, pessoas com dificuldades motoras. Este ponto requer um estudo mais aprofundado das necessidades deste público especial, pois cada caso é um caso único.

8. BIBLIOGRAFIA

8. Bibliografia

- [1] Página oficial da Bioware: Mass Effect 3,
<http://masseffect.bioware.com/agegate/?url=%2F>, visto a 19 de Maio de 2016
- [2] Steam Store: Tom Clancy's EndWar, <http://us.ewo.ubi.com/en/main/game/>, visto a 19 de Maio de 2016
- [3] Página oficial de The Elder Scrolls V: Skyrim,
<http://www.elderscrolls.com/skyrim/>, visto a 19 de Maio de 2016
- [4] Página oficial de Tazti speech recognition software,
<https://www.tazti.com/index.php>, visto a 1 de Junho de 2016
- [5] Página oficial de VoiceBot voice powered game control,
<https://www.voicebot.net/>, visto a 1 de Junho de 2016
- [6] Steam Store: There Came an Echo, <http://store.steampowered.com/app/319740/>, visto a 14 de Abril de 2016
- [7] Steam Store: In Verbis Virtus, <http://store.steampowered.com/app/242840/>, visto a 14 de Abril de 2016
- [8] Kotaku: He believes voice controls can actually work in a video game,
<http://kotaku.com/5991204/he-believes-voice-controls-can-actually-work-in-a-video-game>, visto a 5 de Junho de 2016
- [9] Memeburn: The Pros and cons of the game changing Kinect,
<http://memeburn.com/2010/12/the-pros-and-cons-of-the-game-changing-kinect-review/>, visto a 26 de Junho de 2016
- [10] Asta Speaks: Things to Consider: The Pros and Cons of Voice Recognition Software, <https://astaspeaks.wordpress.com/2013/05/14/things-to-consider-the-pros-and-cons-of-voice-recognition-software/>, visto a 26 de Junho de 2016

- [11] Buzzle: Voice Recognition Software Pros and Cons,
<http://www.buzzle.com/articles/voice-recognition-software-pros-and-cons.html>, visto a 26 de Junho de 2016
- [12] Wikipedia: Speech Recognition,
https://en.wikipedia.org/wiki/Speech_recognition#Applications, visto a 23 de Agosto de 2016
- [13] Wikipedia: Home Automation, https://en.wikipedia.org/wiki/Home_automation, visto a 15 de Julho
- [14] CNET: Talk to your house with these voice-activated smart-home systems,
<http://www.cnet.com/news/talk-to-your-house-with-these-voice-activated-smart-home-systems/>, visto a 15 de Julho
- [15] Nielsen Norman Group: Games User Research: What's different?,
<https://www.nngroup.com/articles/game-user-research/>, visto a 5 de Janeiro de 2017
- [16] Wikipedia: User Experience evaluation: Video Games,
https://en.wikipedia.org/wiki/User_experience_evaluation#Video_games, visto a 14 de Dezembro de 2016
- [17] Game User Research Methods, <http://gameuserr.editme.com/>, visto a 29 de Dezembro de 2016
- [18] Chek Tien Tan, Tuck Wah Leong, Songjia Shen; Combining Think-aloud and Physiological Data to Understand Video Game Experiences; CHI 2014, Toronto, Canada, 381-390, visto a 27 de Dezembro de 2016
- [19] Chek Tien Tan, Tuck Wah Leong, Songjia Shen, Christopher Dubravs, Chen Si; Exploring Gameplay Experiences on Oculus Rift; CHI Play 2015, London, United Kingdom, 253-263, visto a 27 de Dezembro de 2016
- [20] Chek Tien Tan, Sander Bakkes, Yusuf Pisan; Inferring Player Experiences Using Facial Expressions Analysis, IE2014, Newcastle, Australia, visto a 27 de Dezembro de 2016

- [21] Chek Tien Tan, Alessandro Canossa, Pejman Mirza-Babaei, Genevieve Conley, Veronica Zammitto, Günter Wallner; Tool Design Jam: Designing Tools for Games User Research, CHI Play 2015, London, United Kingdom, visto a 27 de Dezembro de 2016
- [22] Chen Si, Yusuf Pisan, Chek Tien Tan; Understanding Players' Map Exploration Styles, ACT Australia, visto a 27 de Dezembro de 2016
- [23] Página oficial da Game Developers Conference, <http://www.gdconf.com/>, visto a 15 de Dezembro de 2016
- [24] Laura Klein; Design for Voice Interfaces, Building Products that Talk; O'Reilly, 2016, visto a 15 de Dezembro de 2016
- [25] Wikipedia: Heuristic Evaluation, https://en.wikipedia.org/wiki/Heuristic_evaluation, visto a 3 de Junho de 2017
- [26] Nielson Norman Group: 10 Usability Heuristics for User Interface Design, <https://www.nngroup.com/articles/ten-usability-heuristics/>, visto a 3 de Junho de 2017
- [27] Melissa A. Federoff; Heuristics and Usability Guidelines for the Creation and Evaluation of Fun in Video Games; Indiana University, December 2002, visto a 3 de Junho de 2017